BEST AVAILABLE COPY

PCT WORLD INTELLECTUAL PROPERTY ORGANIZATION
International Bureau

INTERNATIONAL APPLICATION PUBLISHED UNDER THE PATENT COOPERATION TREATY (PCT)

(51) International Patent Classification 6 : A61B 5/00 | A1

(11) International Publication Number: WO 98/38908

(43) International Publication Date: 11 September 1998 (11.09.98)

(21) International Application Number: PCT/US98/04390

(22) International Filing Date: 26 February 1998 (26.02.98)

(30) Priority Data:
60/039,751 3 March 1997 (03.03.97) US
60/046,809 2 May 1997 (02.05.97) US

(71) Applicant *(for all designated States except US)*: SCHNEIDER MEDICAL TECHNOLOGIES, INC. [US/US]; Suite 200–K, Portola Valley, CA 94028 (US).

(72) Inventor; and
(75) Inventor/Applicant *(for US only)*: SCHNEIDER, M., Bret [US/US]; 140 Balsamina Way, Portola Valley, CA 94028 (US).

(74) Agent: RITTER, Michael, J.; Beyer & Weaver, LLP, P.O. Box 61059, Palo Alto, CA 94306 (US).

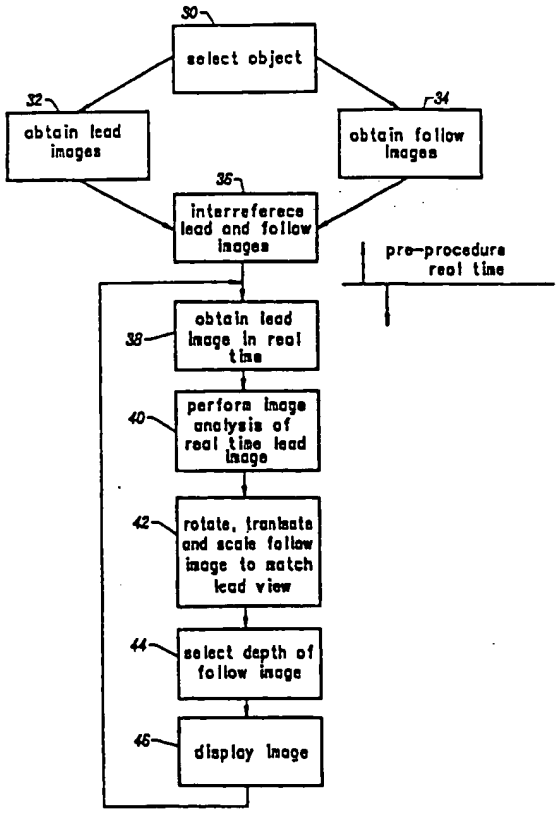
(81) Designated States: CA, JP, US, European patent (AT, BE, CH, DE, DK, ES, FI, FR, GB, GR, IE, IT, LU, MC, NL, PT, SE).

Published
With international search report.
Before the expiration of the time limit for amending the claims and to be republished in the event of the receipt of amendments.

(54) Title: IMAGING DEVICE AND METHOD

(57) Abstract

A method and apparatus for obtaining and displaying a real time image of an object obtained by one modality such that the image corresponds to a line of view established by another modality. In a preferred embodiment, the method comprises the following steps: obtaining a follow image library of the object via a first imaging modality (34); providing a lead image library obtained via the second imaging modality (32); referencing the lead image library to the follow image library (36); obtaining a lead image of the object in real time via the second imaging modality along a lead view (38); comparing the real time image analysis to identify a follow image to correspond to the scale, rotation and position of the lead image (40, 42); and displaying the transformed follow image (46), the comparing, transforming and displaying steps being performed substantially simultaneously with the step of obtaining the lead image in real time.



FOR THE PURPOSES OF INFORMATION ONLY

Codes used to identify States party to the PCT on the front pages of pamphlets publishing international applications under the PCT.

AL	Albania	ES	Spain	LS	Lesotho	SI	Slovenia
AM	Armenia	FI	Finland	LT	Lithuania	SK	Slovakia
AT	Austria	FR	France	LU	Luxembourg	SN	Senegal
AU	Australia	GA	Gabon	LV	Latvia	SZ	Swaziland
AZ	Azerbaijan	GB	United Kingdom	MC	Monaco	TD	Chad
BA	Bosnia and Herzegovina	GE	Georgia	MD	Republic of Moldova	TG	Togo
BB	Barbados	GH	Ghana	MG	Madagascar	TJ	Tajikistan
BE	Belgium	GN	Guinea	MK	The former Yugoslav Republic of Macedonia	TM	Turkmenistan
BF	Burkina Faso	GR	Greece	ML	Mali	TR	Turkey
BG	Bulgaria	HU	Hungary	MN	Mongolia	TT	Trinidad and Tobago
BJ	Benin	IE	Ireland	MR	Mauritania	UA	Ukraine
BR	Brazil	IL	Israel	MW	Malawi	UG	Uganda
BY	Belarus	IS	Iceland	MX	Mexico	US	United States of America
CA	Canada	IT	Italy	NE	Niger	UZ	Uzbekistan
CF	Central African Republic	JP	Japan	NL	Netherlands	VN	Viet Nam
CG	Congo	KE	Kenya	NO	Norway	YU	Yugoslavia
CH	Switzerland	KG	Kyrgyzstan	NZ	New Zealand	ZW	Zimbabwe
CI	Côte d'Ivoire	KP	Democratic People's Republic of Korea	PL	Poland		
CM	Cameroon	KR	Republic of Korea	PT	Portugal		
CN	China	KZ	Kazakstan	RO	Romania		
CU	Cuba	LC	Saint Lucia	RU	Russian Federation		
CZ	Czech Republic	LI	Liechtenstein	SD	Sudan		
DE	Germany	LK	Sri Lanka	SE	Sweden		
DK	Denmark	LR	Liberia	SG	Singapore		
EE	Estonia						

IMAGING DEVICE AND METHOD

BACKGROUND OF THE INVENTION

This invention relates generally to imaging devices and methods and, in particular, to medical imaging devices and methods.

While invasive surgery may have many beneficial effects, it can cause physical and psychological trauma to the patient from which recovery is difficult. A variety of minimally invasive surgical procedures are therefore being developed to minimize trauma to the patient. However, these procedures often require physicians to perform delicate procedures within a patient's body without being able to directly see the area of the patient's body on which they are working. It has therefore become necessary to develop imaging techniques to provide the medical practitioner with information about the interior of the patient's body.

Additionally, a non-surgical or pre-surgical medical evaluation of a patient frequently requires the difficult task of evaluating imaging from several different modalities along with a physical examination. This requires mental integration of numerous data sets from the separate imaging modalities, which are seen only at separate times by the physician. Image-guided surgical systems currently available are vulnerable to line-of-sight obstruction and consequent registration failure. Additionally, the arbitrary orientation of displayed images contributes to confusion and consequent morbidity.

A number of imaging techniques are commonly used today to gather two-, three- and four-dimensional data. These techniques include ultrasound, computerized X-Ray tomography (CT), magnetic resonance imaging (MRI), electric potential tomography (EPT), positron emission tomography (PET), brain electrical activity mapping (BEAM), magnetic resonance angiography (MRA), single photon emission computed tomography (SPECT), magnetoelectro-encephalography (MEG), arterial contrast injection angiography, digital subtraction angiography and fluoroscopy. Each technique has attributes that make it more or less useful for creating certain kinds of images, for imaging a particular part of the patient's body, for demonstrating certain kinds of activity in those body parts and for aiding the surgeon in certain procedures. For example, MRI can be used to generate a three-dimensional representation of a patient's body at a chosen location. Because the physical nature of the MRI imaging apparatus and the time that it

takes to acquire certain kinds of images, however, it cannot conveniently be used in real time during a surgical procedure to show changes in the patient's body or to show the location of surgical instruments that have been placed in the body. Ultrasound images, on the other hand, may be generated in real time using a relatively small probe. The image generated, however, lacks the accuracy and three-dimensional detail provided by other imaging techniques.

Medical imaging systems that utilize multimodality images and/or position-indicating instruments are known in the prior art. Hunton, N., Computer Graphics World (October 1992, pp. 71-72) describes a system that uses an ultrasonic position-indicating probe to reference MRI or CT images to locations on a patient's head. Three or four markers are attached to the patient's scalp prior to the MRI and/or CT scans. The resulting images of the patient's skull and brain and of the markers are stored in a computer's memory. Later, in the operating room, the surgeon calibrates a sonic probe with respect to the markers (and, therefore, with respect to the MRI or CT image) by touching the probe to each of the markers and generating a sonic signal which is picked by four microphones on the operating table. The timing of the signals received by each microphone provides probe position information to the computer. Information regarding probe position for each marker registers the probe with the MRI and/or CT image in the computer's memory. The probe can thereafter be inserted into the patient's brain. Sonic signals from the probe move within the patient's brain. The surgeon can use information of the probe's position to place other medical instruments at desired locations in the patient's brain. Since the probe is specially located with respect to the operating table, one requirement of this system is that the patient's head be kept in the same position with respect to the operating table as well. Movement of the patient's head would require a recalibration of the sonic probe with the markers.

Grimson, W.E.L., et al., "An Automatic Registration Method for Frameless Stereotaxy, Image Guided Surgery, and Enhanced Reality Visualization," IEEE CVPR '94 Proceedings (June 1994, pp. 430-436) discuss a device which registers three-dimensional data with a patient's head on the operating table and calibrates the position of a video camera relative to the patient using distance information derived from a laser rangefinder, cross correlating laser rangefinder data with laser scanline image data with medical image data. The system registers MRI or CT scan images to the patient's skin surface depth data obtained by the laser range scanner, then determines the position and orientation of a video camera relative to the patient by matching video images of the

laser points on an object to reference three-dimensional laser data. The system, as described, does not function at an interactive rate, and hence, the system cannot transform images to reflect the changing point of view of an individual working on the patient. Because the system is dependent upon cumbersome equipment such as laser rangefinders which measure distance to a target, it cannot perform three-dimensional image transformations guided by ordinary intensity images. The article mentions hypothetically using head-mounted displays and positioning a stationary camera "in roughly the viewpoint of the surgeon, i.e., looking over her shoulder." Although the article reminds that "viewer location can be continually tracked," there is no discussion on how the authors would accomplish this.

Kalawasky, R., "The Science of Virtual Reality and Virtual Environments," pp. 315-318 (Addison-Wesley 1993), describes an imaging system that uses a position sending articulated arm integrated with a three-dimensional image processing system such as a CT scan device to provide three-dimensional information about a patient's skull and brain. As in the device described by Hunton, metallic markers are placed on the patient's scalp prior to the CT scan. A computer develops a three-dimensional image of the patient's skull (including the markers) by taking a series of "slices" or planar images at progressive locations, as is common for CT imaging, then interpolating between the slices to build the three-dimensional image. After obtaining the three-dimensional image, the articulated arm can be calibrated by correlating the marker locations with the special position of the arm. So long as the patient's head has not moved since the CT scan, the arm position on the exterior of the patient can be registered with the three-dimensional CT image.

Heilbrun, M.P., "The Evolution and Integration of Microcomputers Used with the Brown-Roberts-Wells (BRW) Image-guided Stereotactice System," (in Kelly, P.J., et al. "Computers in Stereotactice Neurosurgery," p. 196 (Blackwell Scientific Publications 1992)) describe the use of a stereotactic frame with a system for using image analysis to read position markers on each tomographic slice taken by MR or CT, as indicated by the positions of cross-sections of N-shaped markers on the stereotactic frame. While this method is useful for registering previously acquired tomographic data, it does not help to register a surgeon's view to that data. Furthermore, the technique cannot be used without a stereotactic frame.

Goerss, S.J., "An Interactive Stereotactic Operating Suite," and Kall, B.A., "Comprehensive Multimodility Surgical Planning and Interactive Neurosurgery," (both in Kelly,

P.J., et al. "Computers in Stereotactic Neurosurgery, "pp. 67-86, 209-229 (Blackwell Scientific Publications 1992)) describe the Compass™ system of hardware of hardware and software. The system is capable of performing a wide variety of image processing functions including the automatic reading of stereotactic frame fiducial markers, three-dimensional data, and image transformations (scaling, rotating, translating). The system includes an "intramicroscope" through which computer-generated slices of a three-dimensionally reconstructed tumor correlated in location and scale to the surgical trajectory can be seen together with the intramicroscope's magnified view of underlying tissue. Registration of the images is not accomplished by image analysis, however. Furthermore, there is no mention of any means by which a surgeon's instantaneous point of view is followed by appropriate changes in the tomographic display. This method is also dependent upon a stereotactic frame, and any movement of the patient's head would presumably disable the method.

Suetens, P., et al. (in Kelly, P.J., et al. "Computers in Stereotactic Neurosurgery," pp. 252-253 (Blackwell Scientific Publications 1992)) describe the use of a head mounted display with magnetic head trackers that changes the view of a computerized image of a brain with respect to the user's head movements. The system does not, however, provide any means by which information acquired in real time during a surgical procedure can be correlated with previously acquired imaging data.

Roberts, D.W., et al., "Computer Image Display During Frameless Stereotactic Surgery," (in Kelly P.J., et al. "Computers in Stereotactic Neurosurgery," pp. 313-319 (Blackwell Scientific Publications 1992)) describe a system that registers pre-procedure images from CT, MRI and angiographic sources to the actual location of the patient in an operating room through the use of an ultrasonic rangefinder, an array of ultrasonic microphones positioned over the patient, and a plurality of fiducial markers attached to the patient. Ultrasonic "spark gaps" are attached to a surgical microscope so that the position of the surgical microscope with respect to the patient can be determined. Stored MRI, CT and/or angiographic images corresponding to the microscope's focal plane may be displayed.

Kelly, P.J. (in Kelly, P.J., et al. "Computers in Stereotactic Neurosurgery," p. 352 (Blackwell Scientific Publications 1992)) speculates about the future possibility of using magnetic head tracking devices to cause the surgical microscope to follow the surgeon's changing field of

view by following the movement within the established three-dimensional coordinate system. Insufficient information is given to build such a system, however. Furthermore, this method would also be stereotactic frame dependent, and any movement of the patient's head would disable the coordinate correlation.

Drueger, M.W., "The Emperor's New Realities," pp. 18-33, Virtual Reality World (Nov./Dec. 1993) describes generally a system which correlates real time images with stored images. The correlated images, however, are of different objects, and the user's point of view is not tracked.

Finally, Stone, R.J., "A Year in the Life of British Virtual Reality," p. 49-61, Virtual Reality World (Jan./Feb. 1994) discusses the progress of Advanced Robotics Research Limited in developing a system for scanning rooms with a laser rangefinder and processing the data into simple geometric shapes "suitable for matching with a library of priori computer-aided design model primitives." While this method seems to indicate that the group is working toward generally relating two sets of images acquired by different modalities, the article provides no means by which such matching would be accomplished. Nor does there seem to be classification involved at any point. No means are provided for acquiring, processing, and interacting with image sets in real time, and no means are provided for tracking the instantaneous point of view of a user who is performing a procedure, thereby accessing another data set.

As can be appreciated from the prior art, it would be desirable to have an imaging system capable of displaying single modality or multimodality imaging data, in multiple dimensions, in its proper size, rotation, orientation, and position, registered to the instantaneous point of view of a physician examining a patient or performing a procedure on a patient. Furthermore, it would be desirable to do so without the expense, discomfort, and burden of affixing a stereotactic frame to the patient in order to accomplish these goals. Still further, it would be desirable to reduce the vulnerability of the registration process to line-of-sight obstructions. It would also be desirable to utilize such technology for non-medical procedures such as the repair of a device contained within a sealed chassis.

SUMMARY OF THE INVENTION

This invention provides methods and apparatuses for obtaining and displaying in real time an image of an object obtained by one modality such that the image corresponds to a line of view established by another modality. In a preferred embodiment, the method comprises the following steps: obtaining a follow image library of the object via a first imaging modality; providing a lead image library obtained via the second imaging modality; referencing the lead image library to the follow image library; obtaining a lead image of the object in real time via the second imaging modality along a lead view; comparing the real time lead image to lead images in the lead image library via digital image analysis to identify a follow image line of view corresponding to the lead view; transforming the identified follow image to correspond to the scale, rotation and position of the lead image; and displaying the transformed follow image, the comparing, transforming and displaying steps being performed substantially simultaneously with the step of obtaining the lead image in real time.

In another embodiment, the invention provides a method for displaying an image slice of an object comprising the steps of: obtaining a three dimensional follow image of the object; obtaining a real time lead image of the object; transforming the three dimensional follow image to correspond to the lead image; automatically determining a desired depth within the object for observation; generating an image of the object from the transformed three dimensional follow image at the desired depth; and displaying the generated image of the object. The generated image of the object may be a two dimensional image slice.

In another embodiment, the invention provides a method for displaying an image of an object comprising the steps of: obtaining three dimensional follow image data of the object; obtaining a real time lead image of the object utilizing a stereo camera; transforming the three dimensional follow image to correspond to the lead image; and displaying at least a portion of the transformed three dimensional follow image of the object.

The invention is described in further detail below with reference to the drawings.

BRIEF DESCRIPTION OF THE DRAWINGS

The invention, together with further objects and advantages thereof, may best be understood by reference of the following description taken in conjunction with the accompanying drawings in which:

Fig. 1 is a block diagram showing a preferred embodiment of the imaging device of this invention.

Fig. 2 is a flow chart illustrating a preferred embodiment of the method of this invention.

Fig. 3 is a flow chart illustrating an alternative embodiment of the method of this invention.

Fig. 4 shows an embodiment of the invention.

Figs. 5A and 5B show an alternative embodiment of a head mounted display/head mounted camera apparatus.

Fig. 6 shows several methods for determining in real time the depth of slice that may be extracted from a follow image prior to display.

Fig. 7 shows the use of multiple fiducials implanted upon a mobile body part.

Fig. 8 shows a fiducial gun which may be utilized to rapidly and efficiently implant or attach fiducial markers.

Fig. 9 shows an endoscope including dual localizer cameras for acquiring lead images.

Fig. 10 shows a fiducial marker with an elongated staff that penetrates surgical drapes.

Fig. 11 shows a flowchart of an embodiment in which the acquisition of a follow image is controlled in real time.

Fig. 12 illustrates fiducials for positioning a catheter for angioplasty of a blood vessel.

DETAILED DESCRIPTION OF PREFERRED EMBODIMENTS

Definitions.

The following definitions are useful in understanding and using devices and methods of the invention.

Image. As used herein, "image" means the data that represents the special layout of anatomical or functional features of a patient, which may or may not be actually represented in visible, graphical form. In other words, image data sitting in a computer memory, as well as an image appearing on a computer screen, will be referred to as an image or images. Non-limiting examples of images include an MRI image, an angiography image, and the like. When using a video camera as a data acquisition method, an "image" refers to one particular "frame" in the series that is appropriate for processing at that time. Because the ability to "re-slice" a three-dimensional reconstruction of a patient's body in a plane corresponding to the trajectory of the "lead view" (typically the line of view from which the surgeon wishes to view the procedure) is important to this method, the "image" may refer to an appropriately re-sliced image of a three-dimensional image reconstruction, rather than one of the originally acquired two-dimensional files from which the reconstructions may have been obtained. The term image is also used to mean any portion of an image that has been selected, such as a fiducial marker, subobject, or knowledge representation. The term "image" is also intended to encompass any spacially registered data including the receipt of infrared signals in a constant or time-encoded fashion by a CCD matrix. Stereo imager pairs of the same object are herein refered to in the singular as "image."

Imaging modality. As use herein "imaging modality" means the method or mechanism by which an image is obtained, e.g., MRI, CT, video, ultrasound, etc.

Lead View. As used herein "lead view" means the line of view toward the object at any given time. Typically the lead view is the line of view through which the physician, at any given time, wishes to view the procedure. In the case where a see-through head-mounted display and head-mounted camera are utilized, this should be the instantaneous line of view of the physician. As the lead view shifts, all other images should adjust their views to that of the lead view in order to make all of the images that converge to make a resulting composite image accurate.

Lead image. As used herein "lead image" is an image obtained through the same modality as the lead view. For example, if the lead view is the physician's view of the surface of the patient, the lead image could be a corresponding video image of the surface of the patient. Lead images may also be obtained by any real time intraoperative modality, including fluoroscopy, endoscopy, microscopy, ultrasound, or infrared optical localizers. Lead images may be stereo image pairs.

Follow image. As used herein "follow image" will be an image which should be transformed and possibly sliced to the specifications of the lead view and slice depth control. A properly sliced and transformed follow image will usually be in a plane parallel with that of the lead image, and consequently, orthogonal to the lead view, although other slice contours could be used. A properly transformed follow image will be at the same angle of the view as the lead image, but at a depth to be separately determined.

Composite image. As used herein "composite image" is the image that results from the combination of properly registered lead and follow images from two or more sources, each source representing a different modality.

Fiducial marker. As used herein, "fiducial marker" means a feature, set of features, image structure, or subobject present in lead or follow images that can be used for image analysis, matching, coordinate interreferencing or registration of the images and creation of a composite image.

Feature extraction. As used herein "feature extraction" means a method of identification of image components which are important to the image analysis being conducted. These may include boundaries, angles, area, center of mass, central moments, circularity, rectangularity and regional gray-scale intensities in the image being analyzed.

Segmentation. As used herein "segmentation" is the method of dividing an image into areas which have some physical significance in terms of the original scene that the image attempts to portray. For example, segmentation may include the demarcation of a distinct anatomical structure, such as an external auditory meatus, although it may not be actually identified as such until classification. Thus, feature extraction is one method by which an image can be segmented. Additionally, previously segmented areas may be subsequently subjected to feature extraction. Other non-limiting examples of methods of segmentation which are well known in the area of

image analysis include: thresholding, edge detection, Hough transformation, region growing, template matching and the like. See, e.g., Rosenfled, A., "The fuzzy geometry of image subsets," (in Bezdek, J.C., et al., "Fuzzy Models for Pattern Recognition," pp. 340-346 (IEEE 1992)).

Classification. As used herein, "classification" means a step in the imaging method of the invention in which an object is identified as being of a certain type, based on its features. For example, a certain segmented object in an image might be identified by a computer as being an external auditory meatus based on if it falls within predetermined criteria for size, shape, pixel density, and location relative to other segmented objects. In this invention, classification is extended to include the angle, or Cartesian location, from which the object is viewed ("line of view"), for example, an external auditory meatus viewed from 30° North and 2° West of a designated origin. A wide variety of classification techniques are known, including statistical techniques (see, e.g., Davies, E.R., "Machine Vision" Theory, Algorithms, Practicalities," pp. 435-451 (Academic Press 1992)) and fuzzy logic techniques (see, e.g., Bezdek, J.C., et al., "Fuzzy Models for Pattern Recognition," pp. 1-27 (IEEE 1992); Siy, P., et al., "Fuzzy Logic for Handwritten Numeral Character Recognition," (in Bezdek, J.C., et al., "Fuzzy Models for Pattern Recognition," pp. 321-325 (IEEE 1992)). Classification techniques are discussed in Faugeras, "Three-Dimensional Computer-Vision," pp. 483-558 (MIT Press 1989) and Haralick, R.M., et al., "Computer and Robot Vision," vol. 2, pp. 43-185, 289-378, 493-533 (Addison-Wesley 1993).

Transformation. As used herein, "transformation" means processing an image such that it is translated (moved in a translational fashion), rotated (in two or three dimensions), scaled, sheared, warped, placed in perspective or otherwise altered according to specified criteria. See Burger, P., "Interactive Computer Graphics," pp. 173-186 (Addison-Wesley 1989).

Registration. As used herein, "registration" means alignment process by which two images of like to corresponding geometries and of the same set of objects are positioned coincident with each other so that corresponding points of the imaged scene appear in the same position on the registered images.

Description of Preferred Embodiments.

For convenience, preferred embodiments of the invention are discussed in the context of medical applications, such as in brain surgery or other invasive surgeries. The invention is also applicable to other uses, including but not limited to medical examinations, analysis of ancient and often fragile artifacts, airplane luggage, chemical compositions (in the case of nuclear magnetic resonance spectral analysis), the repair of closed pieces of machinery through small access ways, and the like.

The invention improves earlier methods and devices for creating multimodality composite images by providing a new way of selecting and registering the image data. The invention also improves upon earlier methods of image viewing by adjusting to the user's line of sight while in a dynamic filed of view. The user's line of sight (or view) is typically along the path of the eyesight of the user so it is a perspective that is generally related to the physical position of the user. Figure 1 is a block diagram of an imaging system 2 for displaying an image of an object 10 according to a preferred embodiment of this invention. A lead library 12 and a follow library 14 of images of the object 10 obtained by two different modalities communicate with a processing means 16. The imaging modality of either library could be a CT scan, an MRI scan, a sonogram, an angiogram, video or any other imaging technique known in the art. Each library contains image data relating to the object.

Most preferably, at least one of the imaging devices is a device that can view and construct an image of the interior of object 10. The images (or data gleaned from their analysis) are stored within the libraries in an organized and retrievable manner. The libraries may be any suitable means of storing retrievable image data, such as, for example, electronic memory (RAM, ROM, etc.), magnetic memory (magnetic disks or tape), or optical memory (CD-ROM, WORM, etc.).

The processing means 16 interreferences corresponding images in image libraries 12 and 14 to provide a map or table relating images or data in one library to images or data in the other. A preferred interreferencing method is described in detail below. Processing means 16 may be a stand-alone computer such as an SGI Onyx symmetric multiprocessing system workstation with the SGI RealityEngine graphics subsystem (available from Silicon Graphics, Inc.) and suitable software. Additionally, processing means 16 may be an image processor specially designed for this particular application.

A lead imager 18 is provided to obtain an image of object 10 along a chosen perspective or line of view. For example, if object 10 is a patient in an operating room, lead imager 18 may be a video camera that obtains video images of the patient along the line of sight of the attending physician, such as a head-mounted video camera. Preferably, the lead imager is a camera or camera array mounted on the head of the user along his or her line of eyesight. Lead imager 18 sends its lead image to processing means 16 which interreferences the lead image with a corresponding follow image from follow image library 14 and transforms the image to correspond to the lead image. The depth at which the follow image is sliced may be controlled by a depth control 24 (such as a mouse, joy stick, knob, or other means) to identify the depth at which the follow image slice should be taken. The follow image (or, alternatively, a composite image combining the lead image from lead imager 18 and the corresponding transformed follow image from library 14) may be displayed on display 20. Display 20 may be part of processing means 16 or it may be an independent display.

In a preferred embodiment, object 10 has at least one fiducial marker 22. The fiducial marker is either an inherent feature of object 10 (such as a particular bone structure within a patient's body) or a natural or artificial subobject attached to or otherwise associated with object 10. The system and method of this invention use one or more fiducial markers to interreference the lead and follow image or to interreference lead images acquired in real time to lead images or data in the lead image library, as discussed in more detail below.

Figure 2 is a flowchart showing an embodiment of this invention. In the flowchart, steps are divided into those accomplished before the start of the surgical procedure, and those that are accomplished in real time, i.e., during the procedure. In this example, the object of interest is a body or a specific part of the body, such as a patient's head (the follow image modality) and a video image of the surface of the patient's head (the lead image modality). It should be understood, however, that the invention could be used in a variety of environments and applications.

In a preferred embodiment, the lead and follow images are interreferenced prior to the surgical procedure to gather information for use in real time during the surgical procedure. Interreferencing of the lead and follow images gathered in this pre-procedure stage is preferably performed by establishing common physical coordinates between the patient and the video camera

and between the patient and the MRI device. The first step of this preferred method (indicated generally at block 30 of Figure 2) therefore is to mount the patient's head immovably to a holder such as a stereotactic frame.

Next, to gather follow image information, an MRI scan of the patient's head and stereotactic frame is taken, and the three-dimensional data (including coordinate data relating to the patient's head and the stereotactic frame) are processed in a conventional manner and stored in memory, such as in a follow image library, as shown in block 34. The pre-process lead video images of the patient's head are preferably obtained via a camera that automatically obtains digital images at precise locations. Robotic devices built to move instruments automatically between precede stereotactic locations have been described by Young, R.F., et al., "Robot-aided Surgery" and Benabid, A.L., et al., "Computer-driven Robot for Stereotactic Neurosurgery," (in Kelly, P.J., et al., "Computers in Stereotactic Neurosurgery," pp. 320-329, 330-342 (Blackwell Scientific Publications, 1992)). Such devices may be used to move a camera to appropriate lead view angles for the acquisition of the lead library. For example, using the stereotactic frame, the video camera may be moved about the head in three planes, obtaining an image every 2 mm. Each image is stored in a lead image library along with information about the line of view or trajectory from which the image was taken. The stereotactic frame may be removed from the patient's head after all these images have been obtained.

Keeping the patient's head immovably attached to the stereotactic frame during the MRI and video image obtaining steps gives the lead (video) and follow (MRI) image data a common coordinate system. Thus, identification of a line of view showing a portion of a stored video image is equivalent to identification of the corresponding line of view in the stored MRI image. Information interreferencing the stored lead and follow images is itself stored for use for real time imaging during the surgical procedure.

As a final step in the pre-procedure part of the method, the video lead images are digitally analyzed to identify predefined fiducial markers. In a preferred embodiment, the digital representation of each lead image stored in the lead image library is segmented or broken down into subobjects. Segmentation can be achieved by any suitable means known in the art, such as by feature extraction, thresholding, edge detection, Hough transforms, region growing, run-length connectivity analysis, boundary analysis, template matching, etc. A preferred embodiment

of this invention utilizes a Canny edge detection technique, as described in R. Lewis, "Practical Digital Image Processing" (Ellis Horwood, Ltd., 1990). The result of the segmentation process is the division of the video image into subobjects which have defined boundaries, shapes, and positions within the overall image.

The Canny edge detection segmenting technique can be modified depending on whether the image is in two or three dimensions. In this example the image is, of course, a two-dimensional video image. Most segmentation approaches can be adapted for use with either two-dimensional or three-dimensional images, although most written literature concerns two-dimensional image segmentation. One method by which a two-dimensional approach can be adapted for the segmentation of a three-dimensional object is to run the two-dimensional segmentation program on each two-dimensional slice of the series that represents the three-dimensional structure. Subsequent interpolation of each corresponding part of the slices will result in a three-dimensional image containing three-dimensional segmented objects.

The least computationally intensive method of segmentation is the use of thresholding. Pixels above and below a designated value are separated, usually by changing the pixels to a binary state representative of the side of the threshold on which that pixel falls. Using thresholding and related edge detection methods that are well know in the art, and using visually distinctive fiducials, a desired area of the image is separated from other areas. If extracted outlines have discontinuous, simple "linking" algorithms, as are known in the art, may be used to connect closely situated pixels.

If the binarized segmented regions are used for pattern or template matching (between the real time video image and the lead library images), correlations between the video and the follow library are made, according to the methods of the invention. Preferably, the lead and follow images are processed in similar manners, for example by thresholding, so that they can be matched quickly and efficiently. In order to further remove computational load from the processing means, thresholding may be effectively accomplished prior to any processing by the computer by simply setting up uniform lighting conditions and setting the input sensitivity or output level of the video camera to a selected level, such that only the pixels of a certain intensity will remain visible. Hence, only the relevant fiducial shapes will reach the processor. Using methods such as thresholding, with uniform lighting and distinct fiducials, and efficient

classification methods, image analysis as described herein can be accomplished in real time (i.e., at an interactive rate) even using hardware not specially designed for image analysis.

To help resolve the difficulties in segmenting low-contrast points in images (particularly medical images), much effort in the field is being devoted to the development of new segmentation techniques. Particularly likely to be useful in the future are those statistical segmentation techniques that assign to each point a certain degree of probability as to whether or not it is a part of a given segmented object. That probability is based upon a variety of factors including pixel intensity and location with respect to other pixels of given qualities. Once probabilities of each pixel have been determined, assessments can be made of the pixels as a group, and segmentation can be achieved with improved accuracy. Using such techniques, segmentation of a unified three-dimensional file is preferable to performing a segmentation on a series of two-dimensional images, then combining them, since the three-dimensional file provides more points of reference when making a statistic-based segmentation decision. Fuzzy logic techniques may also be used, such as those described by Rosenfeld, A., "The Fuzzy geometry of image subsets," (in Bezdek, J.C., et al., "Fuzzy Models for Pattern Recognition," pp. 340-346 (IEEE Press 1991)).

The final part of this image analysis step is to classify the subobjects. Classification is accomplished by means well known in the art. A wide variety of image classification methods are described in a robust literature, including those based on statistical, fuzzy, relational, and feature-based models. Using a feature-based model, feature extraction is performed on a segmented or unsegmented image. If there is a match between the qualities of the features and those qualities previously assigned in the class definition, the object is classified as being of that type. Class type can describe distinct anatomic structures, and in the case of this invention, distinct anatomic structures as they appear from distinct points of view.

In general, the features of each segmented area of an image are compared with a list of feature criteria that describe a fiducial marker. The fiducial marker is preferably a unique and identifiable feature or set of features on the object, such as surface shapes caused by particular bone or cartilage structures within the patient's body. For example, the system could use an eyeball as a fiducial marker by describing it as a roughly spherical object having a diameter within a certain range of diameters and a pixel intensity within a certain range of intensities. Other potential fiducial markers are the nose, the brow, the pinnae and the external auditory meatus.

Alternatively, the fiducial marker can be added to the object prior to imaging solely for the purpose of providing a unique marker, such as a marker on the scalp. Such a marker would typically be selected to be visible in each imaging modality used. For example, copper sulfate capsules are visible both to MRI and to a video camera. As yet another alternative, the stereotactic frame used in the pre-procedure steps may be left attached to the head. In any case, if an object can be automatically recognized, it can be used as a fiducial marker.

The segmentation, feature extraction and classification steps utilized by this invention may be performed with custom software. Suitable analysis of two-dimensional images may be done with commercially available software such as Global Lab Image, with processing guided by a macro script.

After the images stored in the lead and follow libraries have been interreferenced, and the fiducial markers in the lead images have been identified, the system is ready for use in real time imaging (i.e., images obtained at an interactive rate) during a medical procedure. In this example, real time lead images of the patient's head along the physician's line of sight are obtained through a digital video camera mounted on the physician's head, as in block 38 of Figure 2. Individual video images are obtained via a framegrabber.

In a preferred embodiment, each video image is correlated in real time (i.e., at an interactive rate) with a corresponding image in the lead image library, preferably using the digital image analysis techniques discussed above. Specifically, the lead image is segmented, and the subobjects in the segmented lead image are classified to identify one or more fiducial markers. Each fiducial marker in the real time lead image is matched in position, orientation and size with a corresponding fiducial marker in the lead image library and, thus, to a corresponding position orientation and size in the follow image library via the interreferencing information. The follow image is subsequently translated, rotated in three dimensions, and scaled to match the specifications of the selected lead view. The process of translating and/or rotating and/or scaling the images to match each other is known as transformation. The follow image may be stored, manipulated or displayed as a density matrix of points, or it may be converted to a segmented vector-based image by means well-known in the art, prior to being stored, manipulated or displayed.

Because the follow image in this example is three-dimensional, this matching step yields a three-dimensional volume, only the "surface" of which would ordinarily be visible. The next step in the method is therefore to select the desired depth of the slice one wishes to view. The depth of slice may be selected via a mouse, knob, joystick or other control mechanism. the transformed follow image is then sliced to the designated depth by means known in the art, such as described in Russ, J.C., "The Image Processing Handbook," pp. 393-400 (CRC Press 1992); Burger, P., et al., "Interactive Computer Graphics," pp. 195-235 (Addison-Wesley 1989).

In general, slicing algorithms involves designating a plane of slice in the three-dimensional image and instructing the computer to ignore or to make transparent any data located between the viewer and that plane. Because images are generally represented in memory as arrays, and because the location of each element in the array is mathematically related to the physical space that it represents, a plane of cut can be designated by mathematically identifying those elements of the array that are divided by the plane. The resulting image is a two-dimensional object sliced at the designated plane. Follow images may be displayed according to "perspective rendering" techniques, as are known in the art of computer graphics, so as to most accurately and naturally emulate a given point of view.

In one embodiment, the graphics functions of the system can employ "three-dimensional texture mapping" functions such as those available with the SGI RealityEngine and with the Sun Microsystems Freedom Series graphics subsystems. The SGI RealityEngine hardware/software graphics platform, for example, supports a function called "3-D texture" which enables volumes to be stored in "texture memory." Texel values are defined in a three-dimensional coordinate system, and two-dimensional slices are extracted from this volume by defining a plane intersecting the volume. Thus, the three-dimensional follow image information of this invention may be stored as a texture in texture memory of the RealityEngine and slices obtained as discussed above.

In an alternative embodiment, the three-dimensional data set is held, transformed and sliced in main memory, including in frame buffers and z-buffers, such as those found on the Sun Microsystems 3 graphics subsystem as well as on the Sun Microsystems Freedom Series and SGI RealityEngine graphics subsystems.

The system can display the sliced follow image alone, or as a composite image together with a corresponding lead image, such as by digital addition of the two images. Additionally, the

transformed and sliced follow image can be projected onto a see-through display mounted in front of the physician's eyes so that it is effectively combined with the physician's direct view of the patient. Alternatively, the composite lead and follow images can be displayed on a screen adjacent the patient. The displayed images remain on the screen while a new updated lead image is obtained, and the process starts again.

The imaging system performs the steps of obtaining the lead image and display of the corresponding follow or composite image substantially in real time (or, in other words, at an interactive rate). In other words, the time lag between obtaining the lead image and display of the follow or composite is short enough that the displayed image tracks changes of the lead view substantially in real time. Thus, in the medical context, new images will be processed and displayed at a frequency that enables the physician to receive a steady stream of visual feedback reflecting the movement of the physician, the patient, medical instruments, etc.

In a first alternative embodiment, interreferencing of the images in the lead and follow libraries in the pre-procedure portion of the imaging method is done solely by digital image analysis techniques. Each digitized lead image (for example, a video image) is segmented, and the subobjects are classified to identify fiducial markers. Fiducial markers in the follow images (e.g., surface views of MRI images) are also identified in the same way. A map or table interreferencing the lead and follow images is created by transforming the follow images is created by transforming the follow image fiducial markers. The interreferencing information is stored for use during the real time imaging process. Alternatively, pattern matching techniques may be used to match the images without identifying specific fiducial markers. Davies, E.R., "Machine Vision: Theory, Algorithms, Practicalities," pp. 345-368 (Academic Press 1992); Haralick, R.M., et al., "Computer and Robot Vision," vol. 2, pp. 289-378, 493-533 (Addison-Wesley 1993); Siy, P., et al., "Fuzzy Logic for Handwritten Numeral Character Recognition," in Bezdek, J.C., et al., "Fuzzy Models for Pattern Recognition," pp. 321-325 (IEEE 1992)).

After obtaining the lead and follow image libraries and interreferencing the lead and follow images in the libraries, the method of the first alternative embodiment may then be used to display appropriate slices of the follow images that correspond to lead images obtained in real time. Thus, for example, real time video images of a patient obtained by a video camera mounted on a physician's head can be correlated with lead images in the lead image library via the digital image

analysis techniques described above with respect to a preferred embodiment. The stored interreferencing information can then be used to identify the follow image corresponding to the real time lead image.

The follow image is transformed to match the size, location and orientation of the lead image. The three-dimensional follow image is also sliced to a depth selected via depth control. The transformed and sliced follow image is then displayed alone or as a composite image together with the real time video image. The process repeats when a subsequent real time video image is obtained.

In a second alternative embodiment, the follow images are not sliced in real time. Rather, this embodiment generates a follow image library of pre-sliced follow images obtained on a variety of planes and indexed to multiple lead image lines of view and slice depths. The appropriate follow image slice is retrieved from the follow image library when a given line of view and slice depth is called for by the analysis of the real time lead image. While this embodiment requires greater imaging device memory, it requires less real time processing by the device.

A third alternative embodiment is shown in Figure 3. This alternative embodiment omits the steps of obtaining lead images and interreferencing the lead images with the follow images during the pre-procedure part of the method. Rather, the lead image obtained in real time by the lead imager can be intereferenced directly with the follow images without benefit of a preexisting table or map correlating earlier-obtained lead images with follow images by performing the segmentation and classification steps between the lead image and the follow images in real time or by using other image or pattern matching techniques (such as those described in Haralick, R.M., et al., "Computer and Robot Vision," vol. 2, pp. 289-377 (Addison-Wesley 1993); Siy, P., et al., "Fuzzy Logic for Handwritten Numeral Character Recognition," in Bezdek, J.C., et al., "Fuzzy Models for Pattern Recognition," pp. 321-325 (IEEE 1992)); Davies, E.R., "Machine Vision: Theory, Algorithms, Practicalities," pp. 345-368 (Academic Press 1992)). This third alternative method increases the real time load on the system processor, which could result in a slower display refresh time, i.e., the time between successively displayed images. The slower display refresh time might be acceptable for certain procedures, however. In addition, one advantage of this approach is that it eliminates some of the time spent in the pre-procedure stage.

In another alternative embodiment, the follow images can be obtained in real time and related to the lead images in real time as well. This approach would be useful for use in surgical procedures that alter the patient in some way, thereby making any images obtained prior to the procedure inaccurate.

In other alternative embodiments, the methods shown in Figures 2 and 3 can be practiced using relational data about multiple fiducial markers on the object. For example, instead of determining the orientation of the object by determining the orientation of a single fiducial marker, as in a preferred embodiment, orientation and size information regarding the lead and follow images can be determined via triangulation by determining the relative position of the multiple fiducial markers as seen from a particular line of view. (See "On the Cutting Edge of Technology," pp. 2-14 (Sams Publishing 1993); Moshell, J.M., "A Survey of Virtual Environments," Virtual Reality World Jan/Feb. 1994, pp. 24-36). As another alternative, image analysis techniques can be used to track the movement of the camera or the head rather than its position directly. (See Haralick, R.M., et al., "Computer and Robot Vision," vol. 2, pp. 187-288 (Addison-Wesley 1993); Faugeras, "Three-Dimensional Computer Vision," pp. 245-300 (MIT Press 1989)).

As a further alternative, instead of identifying fiducial markers, pattern matching techniques as described in Davies, Haralick, and Siy may be used for either pre-process or real time matching of corresponding images.

The following is an example of the first preferred embodiment in which the imaging system and method is used to generate and display an image of a patient's head. The two images are: (1) the surgeon's view (produced by a digital video camera mounted on the surgeon's head and pointed at the surface of the patient's head) for the lead image and (2) a three-dimensional CT image of the patient's head as the follow image.

The images are obtained in the pre-procedure stage by a processing computer via a frame-grabber (for the video lead image library) and as a pre-created file including line of view information (for the CT follow image library) and are placed in two separate memory buffers or image libraries. As previously described, the lead images and follow images are preferably obtained while the patient wears a stereotactic head frame. Using the frame's precision instrument guides (preferably, but not necessarily, with a robotic device), numerous video images are taken

from a variety of perspectives around the head. Each image is stored in the lead image library along with the line of view, or trajectory, along which that image was obtained. The stereotactic frame is then removed.

The images in the lead image library are interreferenced with images in the follow image library by correlating the lines of view derived in the image obtaining steps. This interreferencing information is used later in the real time portion of the imaging process.

After gathering the pre-procedure lead and follow image information, the imaging system may be used to obtain and display real time images of the patient. In this example, the real time lead image is obtained via a head-mounted video camera that tracks the physician's line of sight. Each real time lead video image is captured by a frame grabber and analyzed to identify predetermined fiducial markers according to the following process.

The real time lead images are segmented via the Canny edge detection technique (Lewis, R. "Practical Digital Image Processing," pp. 211-217 (Ellis Horwood Limited (1990)), which identifies the boundaries between different structures that appear in an image. The fiducial marker for this example is the eye orbit of the patient's skull, which has been enhanced by drawing a circumferential ring with a marker pen. The orbital rims can been seen both on the surface of the face with a video camera as bony ridges. To perform the classification step, the computer might be told, for example, that a left eye orbit is a roughly circular segmented object with a size between the threshold numbers of 0 and 75, which occurs on the left side of the video images.

From various angles of view, the orbits appear as ellipses, once they have been segmented. When viewed face-to-face with the patient, the ellipses representing the orbits will, at least when considered as a pair, most closely approximate circles. In mathematical/image analysis terms, that is to say that the major axis (the long axis of an ellipse) is most closely equal to the minor axis (the short axis of an ellipse). As one moves along the x axis, the horizontal axis becomes increasingly shortened, lowering the "axis ratio." At the same time, the "ellipse angle" (the angle in degrees between the major axis and the x axis) is approximately 90°. By contrast, as one moves along the y axis, the axis ratio of the ellipses also decreases accordingly, but the ellipse angle is now approximately 0°.

One can appreciate that any combination between these extremes of pure vertical and pure horizontal viewpoint changes would be accordingly reflected in the axis ratio and ellipse angle

measurements. Hence, any given view can be determined, or classified, as being along a certain line of view. Left and right views will not be confused because of the spatial relationship between the two ellipses and other fiducials (one orbit is to the left of the other relative to some other (third) fiducial). In this way, a computer program can be "taught" that an ellipse of given shapes and orientation correspond to the head at a specific orientation. Major and minor axes and their ratio are calculated by well-known formulas (Pratt, W.K., "Digital Image Processing," p. 644, (John Wiley & Sons 1991)), and are a standard feature in commercially available software packages like Global Lab. Such tools also make it possible to analyze images so that they can be "matched" to other images which show the fiducial markers from the same perspective. Alternatively, if a mask-shaped image that includes both orbits and the nose bridge is extracted morphologically, it will also have an unambiguous shape.

After the orbits have been identified, the derived orientation of the real time lead image is compared to the stored information regarding the pre-procedure lead images to identify the pre-procedure lead image that corresponds to the physician's line of view. Because of the earlier interreferencing of the lead and follow images, identification of the lead image line of view will provide the correct follow image line of view. If the real time line of view does not correspond exactly with any of the stored lead image lines of view, the system will interpolate to approximate the correct line of view.

After determination of the correct line of view, the follow image must be translated, rotated and scaled to match the real time image. As with the line of view, these transformation steps are performed by comparing the location, orientation and size of the fiducial marker (in this example, the orbit) of the real time video image with the same parameters of the fiducial marker in the corresponding lead library image, and applying them to the follow image, in combination with a predesignated scaling factor which relates the size of the images in the lead and follow libraries. Of course, any standard or arbitrarily selected position, orientation (e.g., axial, coronal, saggital) or scale may be viewed.

After any transformation of the follow image, the follow image must be sliced at the appropriate depth. The depth can be selected by use of an input mechanism associated with the system, such as a mouse, knob, joystick, switch on a hand-held probe, or keyboard. The resulting follow image slice is then displayed on a head-mounted, see-through display worn by the

physician, such as the displays marketed by RPI Advanced Technology Group (San Francisco, CA) and by Virtual Reality, Inc. (Pleasantville, NY). The process repeats either on demand or automatically as new real time lead images are obtained by the video camera.

Stereoscopic displays can be a useful way of displaying follow images or composite images to give a three-dimensional appearance to the flat displays of CRT's and head-mounted displays. Stereoscopic displays can also improve the effectiveness of the invention by giving appropriate depth cues to a surgeon. The haed mounted display/camera appartus may include surgical loupes for magnification and/or lights for improved illumination of the surgical field.

In the context of the current invention, various methods of producing a three-dimensional view to a user may be used with relative ease. In one embodiment, a head-mounted camera is fixed very close to the user's non-dominant eye; the parallax between the user's natural ocular view and the synthetic view displayed on the see-through head-mounted display creates an approximation of the correct three-dimensional view of the image.

In another embodiment, alternating polarized light filters such as those in the Stereoscopic Display Kits by Tektronix Display Products (Beaverton, OR) between the user's eyes and a stereoscopic display are used. The stereoscopic system displays artificially parallaxed image pairs which provide a synthetic three-dimensional view. Such stereoscopic views are produced and displayed by means well known in the art and may be displayed on any display device, including a conventional CRT or a see-through head-mounted display. This method provides the user, such as a surgeon, with a very precise illusion of seeing the exact three-dimensional location of a specific structure within a patient's body. Such a method not only provides increased realism to the images provided by the invention, but also helps make image guided surgical procedure more accurate, safe and effective.

The speed and efficiency of the hardware used with this invention may be improved by the use of specialized subsystems, leaving the full power of the host system available for miscellaneous tasks such as communicating between the subsystems. Thus, for example, while the Onyx workstation can be used for all vision processing tasks, specialized machine vision subsystems, such as the MaxVideo 200 and the Max860 systems (Datacube, Inc., Danvers, MA) or the Cognex 4400 image processing board (Cognex Corp., Needham, MA), may be used together with the Onyx. These subsystems are designed to take over from their host system

computationally intensive tasks such as real-time edge detection, extraction of shapes, segmentation and image classification.

In another configuration, MaxVideo 200 and Max860 subsystems reside on VME busses of an Onyx with a RealityEngine, with all subsystems under control of the Onyx. In another configuration, MaxVideo 200 and Max860 subsystems are under the control of SPARC LXE (Themis Computer, Pleasanton, CA) all residing on VME busses of an Onyx with a RealityEngine. In another configuration, MaxVideo 200 and Max860 subsystems reside on a SPARC 20 workstation with a Freedom Series 3300 graphic subsystem (Sun Microsystems, Mountain View, CA), which has z-buffers and tri-linear MIP texture mapping features. In yet another configuration, MaxVideo 200 and Max860 subsystems reside on a SPARC 20 workstation with an SX graphics subsystem (Sun Microsystems, Mountain View, CA). In any of the above cases, the MaxVideo 200 subsystem performs integer-based image processing, filtering, image segmentation, geometric operations and feature extraction, and image classification (lead image derived transformation instructions) and evaluation tasks, communicating its computational output, directly or indirectly, to the graphic subsystem. The Max 860 subsystem may be used to perform similar functions, if desired, which require floating point calculations.

Also, a variety of operating systems can be used, depending upon what hardware configuration is selected. These operating systems include IRIX (Silicon Graphics, Inc., Mountain View, CA), SunOS/Solaris (Sun Microsystems, Inc.), or VXWorks (Wind River Systems, Inc., Alameda, CA).

The invention can be used as part of the vision system of remote-controlled machines (such as remote-controlled military vehicles) and autonomous robots (such as surgical robots). Follow image views or composite views generated according to the method of this invention may be used for guidance through an area that is obscured to the view of the naked eye or video camera but known by some other means. For example, if the exterior of a building is visible, and a CAD-type model of that building is also available, a military device can target any room within that building based upon the exterior view. Appropriate follow images or composite views may be used directly in the autonomous vision systems of robots by means well known in the robotics art or may be used by a remote or local human operator.

Modifications are possible without departing from the scope of this invention. For example, the imaging modalities could be angiography (done preoperatively) and fluoroscopy (done in real time and used as either a lead or follow image), so that the location of a medical instrument inserted into a patient's body can be tracked in real time.

Fluoroscopy may be used as a method of read time intra operative lead image acquisition. Natural or artificially implanted fiducial markers on the surface of the patient, on a catheter tip, and/or affixed within deep tissue may be identified by the computer system as long as they are visible to a fluoroscopic camera. Examples of this kind of fiducial marker placement are shown in Figure 12.

Figure 12 illustrates fiducials for positioning a catheter for angioplasty of a blood vessel. A catheter 450 is shown within the lumen blood vessel 452 which is shown with plaque. At the tip of the catheter is a fiducial 454 and there is another fiducial 456 on the catheter so that the fiducials 454 and 456 may be utilized to delineate balloon 458 for performing angioplasty. A fiducial 460 is shown attached to the exterior of blood vessel 452 with barbs. Additionally, a fiducial 462 is shown attached to the exterior surface of skin 464. Fiducials 460 and 462 may be attached any number of ways including barbs, small sutures, adhesive material, and the like. The fiducials allow the 3-dimensional position of anatomical structures to be derived in real time.

The positional information may be uesed to dictate instructions for follow image acquisition or transformation as further described in this document, and the placement of instrument effigies in the appropriate location. In the case of endovascular approaches to surgery, for example, fiducial markers on a catheter tip and fiducial markers on or in a patient's body adjacent to the blood vessel in need of treatment may both be tracked using fluoroscopic lead image acquisition, and obtaining from this 2-dimensional data appropriately acquired or transformed and edited 3-dimensional contrast CT data. When an endoscope tip has a fluoroscopically trackable fiducial marker, the desired view of the tip of the scope may be obtained accordingly, and corresponding follow image views may be displayed by the system.

Furthermore, although the examples described above primarily use single body markers (e.g., eyes, ears) as the key to establishing a line of view, it is anticipated that the simultaneous consideration of many features and the determination of a best match during classification would yield the most accurate at determining source image orientation the computer will become.

Furthermore, by considering more features in the object being recognized, additional source image data can be obtained. For example, the area of the ellipses can be used to correlate the sizes of the two images during the scaling process. Artificial markers, such as foil of various shapes pasted on the skin, clothing, or surgical drapes may also serve the same purpose.

Fiducial markers for image guided surgery are commercially available from several sources including bone-implantable fiducials made by ACT Medical (Newton, MA), and adhesive skin markers made by E-Z-EM (Westbury, NY).

Figure 10 shows a fiducial marker with an elongated staff that penetrates surgical drapes and may be detected by a localizer camera. A fiducial 350 includes a fiducial array 353 at one end and barbs 354 at the other. The fiducial marker is shown penetrating a surgical drape 356, skin 358, and being affixed to internal organ 360 such as the liver. As shown, the fiducial marker it affixed to the internal organ by extensible and retractable tenaculum barbs, but other mechanisms may be utilized.

Fiducial markers such as fiducial marker 350, whether adhered to the skin's surface, implanted within bone, or implanted in deep tissue, may have elongated shafts between the affixed base and the camera localizer visable surface. Such a shaft allows the marker to pass through surgical drapes and even layers of tissue. Consequently, a long-shafted fiducial marker can transcend any physical obstructions lying between an organ that needs to be tracked and a surface visible to localizer camers. In one example, fiducial markers are adhered to the patitent's scalp in a conventional manner. However, the long shaft penetrates through the surgical drapes that may be taped about the shaft for more sterile coverage.

In another embodiment, under ultrasonic guidance so as to avoid critical structures, a fiducial marker may be placed through the skin on a patient's skin into an internal organ such as a liver or prostate. Subsequently, the patient may be MRI or CT scanned. Despite the fiducial being anchored in the liver, and moving about with any internal shifting of that organ, the long shaft that penetrates to the surface is visible to the lead image localizer cameras.

It is possible to use more than two different imaging modalities to prepare a composite image, with one of the images serving as a "linking" image for the purpose of matching fiducial markers in the other two images. For example, the anterior commissure and posterior commissure of the brain might be visible on both MRI and CT. Hence, those common points of reference

allow two entirely separate image coordinate systems to be related to one another. Hence, the "follow image" could be a composite of data obtained by several modalities, previously registered by established means (Kelly, p. 209-225), or a series of separate follow images sequentially registered to each other, or to the lead image by methods herein described. In this way, a surface video camera could be correlated with the CT via the MR coordinate link.

In yet another embodiment, a surgical instrument may be tracked using the techniques of this invention and displayed along with the lead and/or follow images. For example, images of an instrument may be obtained using a video camera or fluoroscopy. If the dimensions of the instrument are known, the image of the instrument may be related to three-dimensional space and displayed with respect to the lead and/or follow images of the patient, even if part of the instrument actually cannot be viewed by the video camera or fluoroscope. This is possible because, like fiducial body features, instruments generally have unique appearances which are characteristic points from which they are viewed. While tracking instruments, a real-time imaging modality could be used as either a lead or a follow image. Because instrument movement may occur independently of the position of the physician and the patient, instrument tracking tasks are preferably performed independent of a patient tracking system, such as by a separate computer or separate processor running in parallel with the computer or processor tracking the patient. Both computers may, of course, derive their input form the same video lead images, and their displays are preferably composited into a single unified display. Alternatively, instruments may be tracked by electromagnetic, sonic or mechanical motion detector systems known in the art. Some such methods are discussed by Kelly, P.J., et al., "Computers in Stereotactic Neurosurgery," pp. 353-354 (Blackwell Scientific Publications, 1992)). Such instruments may bear additional features such as knobs, buttons or switches for controlling image acquistion or display.

The goal of the image guidance in such a case might be a task such as automatic robotic positioning. For such a purpose, a user's line of sight may be used to cue the activation of stepper motors that, for example, move a robotic arm or robotic camera to a specific position or orientation. Further extension of this principle would include the case of the "user" of the invention being a semiautonomous machine, whose computerized "view" is important to the execution of a specific task.

Alternatively, a localizer camera or camera pair may be placed on or within an instrument such as the head of a surgical microscope, or upon an endoscope as shown in Figure 9. Figure 9 shows an endoscope including dual localizer cameras for acquiring lead images. An endoscope 300 includes an endoscope shaft 302 and localizer cameras 304 secured to the shaft. The fiducial markers identified by the system in order to localize the instrument need not be within the optical field of view of such a microscope or endoscope. In other words, the lead image is processed by the computer system and need not be the same as the view that is provided to the eyes of the user. Endoscope-mounted cameras, like head mounted cameras, are freely movable in three dimensions wihtout motoric or mechanical assistance, and require the active effort of the user in order to maintain line-of-sight.

One problem encountered by neurosurgeons who use operating microscopes in conjunction with overlaid images derived from sources such as MRI or CT is that the slice depth displayed as an overlay may not properly reflect the exposed anatomical surface upon which the surgeon is working. Thus, getting the image sliced to correspond with a physical object being viewed is an area in need of accurate automation. Automated slice-depth selection may be accomplished by making the distance of the user's head/lead image acquisition device directly correlated with the z-value (or equivalent) based slice section of a three dimensional image. In such a manner, the computed selection of the coordinates to be rendered transparent, or otherwise graphically ignored in the display, and those that are to be rendered opaque or otherwise visible, may be a function of how far the user's perspective is from a given target.

In practical terms, this may be accomplished with a variety of distance-measuring devices that are known in the art. Such devices run a wide gamut from mechanical arms with potentiometers of optical encoders in their joints, to sound or electromagnetic energy based technologies. Any technology capable of quickly and accurately determining the precise distance between a visual target and a predetermined point may be used for this purpose.

For example, a commercially available range finder, for example a laser range finder, may be used for this purpose. In such an embodiment, a pulsed signal is emitted from a source that may be located on a surgeon's head mounted display, or on the optical head of an operating microscope. As an alternative to conventional rangefinders, a displacement meter, such as that manufactured by Keyence Corp. (Woodcliff Lake, NJ) may be used for precisely ascertaining the

distance to a visual target. Using a displacement meter, a laser beam emitter, for example, sits near, and in a precise orientation with respect to a CCD camera array, both facing the target. The precise position location on the CCD array that is activated in response to the laser striking the target surface can be computed to reveal the precise distance to target. One may use a commercially available displacement meter, or may use the position of laser beam on the lead image to similarly derive the distance measurements as part of the processing of the lead image. In such a case, the lead image would not only provide scale, translatory position and rotation instructions, for a follow image, but also depth slicing instructions.

Image processing techniques may also be used to determine distance to target. For example, images captured by CCD cameras are often automatically focused using a variety of image processing algorithms that, for example seek to minimize the width of strong boundary lines between objects in the image. By measuring the degree of such processing that must be carried out to bring the image into focus, one has a measure of distance. Depth of slice to be displayed may also be determined by means such as the maximal depth of the tip of a tracked instrument, once it is beneath the surface of the patient's skin.

Once the distance between, for example, the head of an operating microscope and a specific exposed anatomical structure is measured, one may register that particular distance to a given slice depth in the follow image set. For example, the distance between the head of a operating microscope and the surface of a patient's scalp at the occupant may be registered with respect to a that same point on in an MRI data set. As an operation proceeds, and the patient's skull and brain is incised, deeper surfaces will be exposed, and these deeper levels are reflected as longer distances by the range finding device. The measured longer distance is reflected, in turn, by a correspondingly deeper selected slice depth for the image overlay to be displayed. Of course, the slice depth/distance relationship may be recalibrated at any time, such as when the resting distance between the optical head of the microscope and the patient is changed. A second range finding or position assessing device may be used to reach a known marker and hence the new position of the optical head, and provide this information to the processing means in order to ascertain a corrected distance to target with which to choose an appropriate slice depth.

In an alternative embodiment, this system may dictate the size, position, rotation slice depth, etc., of a follow image before or during its acquisition. Such an embodiment is of

particular value in the case of follow images being acquired in real time. In such an embodiment, the transformation instructions provided by the machine vision subsystem herein described are relayed to the computerized image acquisition controls or MRI or CT machine, rather than to a graphics manipulation subsystem. Computerized image acquisition controls are a standard part of modern medical imaging equipment, including CT and MRI machines, such as those produced by GE Medical Systems, and by Siemens. These computerized image acquisition controls typically use manual entry, such as by a keyboard or mouse to select the scale, rotation, translatory position, and slice depth of images to be acquired, typically using a "localizer" image as a reference. In this embodiment, however, the keyboard and mouse are bypassed, and the scale, rotation, translatory position, and slice depth instruction set are automatically communicated to the computerized image acquisition controls by the machine vision subsystem. This instruction set may is essentially the same as the transformation instruction set described in previous embodiments, as both simply reflect the rotation, translatory position, and/or scale of the lead image, plus the slice depth instruction from the slice depth control. In this manner, the follow image may simply be acquired in real time in its desired translation, rotation, scale and/or slice, rather than being transformed into that desired form from a previous form.

Embodiments in which data derived from lead image localization is used to control the parameters of a follow image acquisition are particularly useful for follow image acquisition apparatuses such as "open MRI" machines and "intraoperative CT" machines. Such follow image acquisition devices allow efficient access to the patient while scans are in process. Figure 11 outlines the process by which lead images can be used to dictate the orientation parameters of a concurrent MRI or CT scanning process.

Figure 11 shows a flowchart of an embodiment in which the acquisition of a follow image is controlled in real time. At a step 400, the relative 3-dimensional location and orientation of objects is ascertained. The system may then calculate the scan acquisition parameters at a step 402. The scan acquisition parameters may include such information as the depth of a slice that is desired to be acquired as the follow image.

At a step 404, the system translates the scan acquisition parameters into instructions for directing the real time scanning device to acquire the desired follow image. The follow image may be displayed at a step 406.

As previously discussed, the characteristics of an object as seen within a lead image by a machine vision subsystem of this invention may show a number of parameters including, but not limited scale, rotation, translatory position. As herein described, slice depth may be designated by both manual and by automatic means. Never the less, it is not essential that all of these instructions be applied to the follow image in order to gain the benefits of this invention.

For example, when the invention herein described is applied to an operating microscope, the scale of the image shown as an overlay to the optical view of the microscope may be at a magnification factor of the lead image, or at a size completely unrelated to that of the lead image. In such a case, one may choose to determine only the translatory and/or rotational position of an image overlay based upon lead image instructions. Alternatively, for example the case of MR follow images acquired in real time, one may wish to transform the acquired image by none, some, or all of these parameters, the remaining properties being inherent to the acquired image as dictated by the acquisition instruction set.

Alterations in the external and internal anatomy occur during surgical procedures. Ideally, these alterations are reflected in high-quality real time follow images that are continuously acquired throughout the procedure. In many cases, though, real time follow imaging is not available, or has inadequate ability to clearly show the ongoing anatomical changes. In such cases, one may wish to perform online image editing of a previously acquired follow image, in accordance with specific movements of specific surgical instruments being tracked. One example has been previously discussed herein: tracking instruments within the body by superimposing computer generated graphic effigy of that instrument, in its property orientation, over the follow image being displayed. Taking this methodology further, one may erode the pixels of a given portion of an image by a predesignated amount when they are have been touched by an effigy of an activated eletrocaudery instrument for a predesigned amount of time. Computationally, this is accomplished by tracking the position of an instrument as previously described, and ascribing a graphical eroding behavior to certain graphical locations on the virtual instrument. When these locations coincide with image locations of predesignated pixel values (representing certain types of tissues), an image processing routine such "erode" (a standard routine well known in the art of image processing) is initiated at that location. This methodology may be extended, for example, to include the graphical shrinking abscess in response to an automatically measured volume of

aspirated fluid. As another example, a portion of an image may be divided in response to the movement of a cutting tool in the region that the image represents. Of course, manual editing of the follow image data set may also be accomplished in accordance with the wishes of a user, employing manual image editing methods and software that are well known in the art.

As an alternative embodiment, intraoperatively acquired images may be used to control the automated intraoperative editing of follow images that have been previously acquired by another modality. Such real time images may be lead images, or may be independently acquired. In one such embodiment, changes between an image taken by an endoscope from a designated location, and another image taken at a later intraoperative time are digitally ascertained, and these changes are then mapped onto the previously acquired follow image as an edit. Thus, even a previously acquired follow image can be modified to simulate one created in real time. In another embodiment, real time ultrasound may be mapped onto, for example, a previously acquired MR image. In yet another embodiment, the temporal changes in the lead image view as provided by a head mounted camera or by an operating in the lead view as provided by a head mounted camera or by an operating microscope optical head may be used to modify a previously acquired follow image.

Many techniques in the emerging field of computer based tissue modeling are expected to be used within the methods herein described. For example, the manner in which tissue deforms in response to a probe pushing against it may be modeled in terms of tissue elasticity and other parameters, so as to make intraoperative follow image edits as accurate as possible.

In order to have good spatial registration of a follow image, the viewpoint of the lead image acquisition device should very closely approximate the actual point of view of the user. Ideally these two views would be identical. One way of making the two views identical is to use a beam splitter. A beam splitter is an optical device, well known in the art, which essentially takes light that enters from one side, and divides it equally in two different directions. Hence, anyone looking at either of the two beam splitter outputs would see the same thing. Placing a beam splitter on a see-through head mounted display, for example, one may channel the same view that he sees with his eyes, to a CCD camera or other lead image acquisition device. This ensures that the machine vision subsystem is seeing the same thing as the user is, thus permitting the most accurate registration and compositing of a follow image in the user's visual field.

In an alternative embodiment shown in Figure 4, the hardware for the system may consist of specialized subsystem boards within the chassis of an IBM-compatible PC, such as a dual Pentium Pro system (Intel Corp., Santa Clara, CA). The hard drive and memory should each be at least large enough to hold all data sets being processed, all programs being run, and all correlation information relating lead images to follow image transformation, acquisition, and/or slicing instructions. For example, an Octree (Octree Corporation, Curpertino, CA) 3D graphics board (PCI) and Cognex Corporation (Mountain View, CA) 5600 machine vision board (ISA), both in a PC. The use of specialized subsystem boards on a PC platform optimizes the price/performance ratio of the system. Head-mounted displays, see-through or non-see-through, such as those made by Kaiser Electro-optics (Carlsbad, CA), or by Virtual I/O (Seattle, WA) are well suited to the display purpose, as are digital CCD endoscopes as are known in the industry and operating microscopes such as those made by Carl Zeiss, Inc. (Thornwood, N.Y).

Translucent volume renderings can show multiple depths of a given follow image at once. Volume rendering is a method of medical image display that is well known in the art, and may be accomplished by processing techniques such as ray casting. Volume renderings can be done using Octree hardware and/or software, as well as other commercially available volume rendering hardware and/or software, and may be done in conjunction with slicing. Because volume rendering shows 3 dimensions of data before the eye at once, it can reduce the degree of precision necessary for selecting a slice depth for display.

Figures 5A and 5B show an alternative embodiment of a head mounted display/head mounted camera apparatus. In this particular embodiment an immersive (non-see-through) head mounted display is worn in the "semi-immersive" position (i.e., high enough that the user can see an unobstructed view when looking in the lower margin of his visual field, but sees the display when looking in the upper margin of his visual field). This allows a surgeon, for example, to work on an operation with the unobstructed view to which he is accustomed, but giving the surgeon the ability to see the pertinent computer data by simply glancing upward slightly. In such a case, the follow image, or surface/follow image composite display may be offset slightly in the vertical plane from its actual location, so as to provide a view corresponding to that which is seen through the unobstructed path below. Note that stereo CCD cameras 110 and 112 are shown at the level of the user's eyes, not the level of the display.

Also shown in figures 5A and 5B is the use of a stereo camera pair for acquiring lead images. In such an embodiment, the lead image used is actually a stereo image pair; the difference in perspective between the two images helps to discern spatial differences that are more difficult to discern with a single image, by means well known in the art.

The lead imager acquisition and analysis subsystem may be a commercially available optical tracking system such as the field matrix CCD-based Polaris system by Northern Digital, Inc. (Waterloo, Ontario, Canada), or the linear CCD-based tracking systems by Image Guided Technologies (Boulder, Colorado). In the case of the Polaris system, the localizer camera used takes the form of a stereo camera pair so as to improve the accuracy of localization. Left and Right CCD matrices may be placed lateral to each eye, so as to maximally emulate the user's visual perspective without obstructing the user's eyesight, an examplary embodiment is shown in Figures 5A and 5B. Display means may also include "virtual retinal display" technologies that are known in the art.

One advantage of placing localizers along the line of sight of the user's eyes is that the machine vision subsystem's view of the fiducials is not likely to be obstructed without also obstructing the eysight of the user. Another advantage is the ease with which this enables a user-centric display to be rendered.

Surface images may be displayed alone or as part of a surface/depth composite on head mounted display 105, and may be derived from stereo images from stereo lead cameras 110 and 112, or may be derived from a third, centrally located camera 125. Camera 125 provides a monocular image that approximates the view provided by the two eyes of a user.

Figure 6 shows several methods for determining in real time the depth of slice that may be extracted from a follow image prior to display. An optically tracked probe 130 including standard fiducials 135, but for this example could also be tracked by any other standard method including position-sensing arms, and other methods known in the art. In such an embodiment slice depth of the follow image may be cued by the position of the probe in space. For example, the maximum depth of the probe (surface 190) within a patient's body, as computed in real time may be used as the criteria for slice depth selection.

Figure 6 also shows the use of a light emitter (in this case a laser)/detector pair 140. The time required for a pulse emitted by laser 150 to return to the adjacent detector 145 is a direct

function of distance to target surface 190. Hence, if the laser is aimed at the bottom of a surgeon's excavation of a body part (or at any structure of interest, for that matter), that structure will be exposed in the subsequently sliced follow image. In this manner, depth of excavation can be tracked in real time, and accordingly reflected in the manner in which the follow image is displayed.

Figure 6 shows CCD camera 160 which is aimed at target surface 190, for example the bottom of a surgical excavation. Automatic focusing algorithms required to bring the image of the target into maximum sharpness can be used to calculate the distance to target, when supplied with optical characteristics of the lens.

Referring still to Figure 6 a depth finding device 170 is shown in which two laser beams project toward target surface 190. Laser tube 172 to sits fixed in angle with respect to the device, while laser tube 174 is adjustable with respect to the device. Adjustments of the trajectory of laser tube 172 may be accomplished by means known in the art such as threaded knob 178, and the trajectory may be monitored by a variety of means know including position trackers such as linear potentiometer 176. Consequently, the relative positions of beam emissions are known by a computer that monitors device 170 via cable 180. If device 170 faces a target surface 190 (e.g., the bottom of a surgical excavation, or a point of interest), and laser tube 174 is adjusted so as to make the beams form laser tube 174 and 172 converge into a single point, that point of convergence corresponds to a specific, precise distance form device 170. In this manner, precise depth of slice required of a follow image may be dictated by the depth finding device 170 in real time.

Devices such as 130,140, 160, and 176 may be mounted at a fixed location by a positionable arm 181 above the surgical field prior to surgery, and registered to that location by means known in the art. Fixation, may occur, for example via a clamp 182. Alternatively, depth measurement devices may be tracked in real time by mean described herein, as well as other means known in the art, and their output may be interpreted by the computer to adjust for their real-time spatial location.

Figure 7 shows the use of multiple fiducials implanted upon a mobile body part (in this case intestine 205 within abdomen 200). With multiple fiducials 210, 215, and 220 in place, even mobile body parts can be tracked in real time as they move about. By making each fiducial

individually recognizable and distinct from the others, the three dimensional attitude of that mobile structure may be reconstructed, either by doing online modification of the follow image. Means for accomplishing this are known in the art, including those techniques used to modify cartoon images on a computer screen in accordance with the motions of a live actor. Such techniques are accordingly readily adaptable to the monitoring of mobile medical structures. Also note that fiducial 220 is asymmetric in shape. The polarity of an asymmetrical fiducial marker can help to provide 3D-orientation information from a single fiducial marker.

Figure 8 shows fiducial gun 250, which can be used for rapidly and efficiently implanting or attaching fiducial markers 260 and 290 onto surfaces. The fiducial markers 260 and 290 may be held in place my means known in the art including retention prongs 270. The device operates in a manner similar to surgical staple and clamp guns that are known in the art.

All references cited herein are incorporated herein by reference in their entirety. The instant invention is shown and described herein in what are considered to be the most practical and preferred embodiments. It is recognized, however, that departures can be made therefrom which are within the scope of the invention, and that modifications will occur to those of skill in the art upon reading this disclosure.

CLAIMS

1. A method for displaying an image slice of an object, comprising the steps of:

obtaining a three dimensional follow image of the object;

obtaining a real time lead image of the object;

transforming the three dimensional follow image to correspond to the lead image;

automatically determining a desired depth within the object for observation;

generating an image of the object from the transformed three dimensional follow image at the desired depth; and

displaying the generated image of the object.

2. The method of claim 1, wherein the generated image of the object is a two dimensional slice at the desired depth.

3. The method of claim 1, wherein the desired depth is an exposed anaotomical surface during surgery.

4. The method of claim 1, wherein the desired depth is automatically determined utilizing a location of an optically tracked probe.

5. The method of claim 1, wherein the desired depth is automatically determined by measuring time for light to be emitted and return.

6. The method of claim 1, wherein the desired depth is automatically determined by an focusing algorithm of a camera.

7. The method of claim 1, wherein the desired depth is automatically determined by a range finder.

8. The method of claim 1, wherein the desired depth is automatically determined by utilizing a convergence of two laser beams.

9. The method of claim 1, wherein the generated image of the object is displayed on a head mounted display.

10. The method of claim 9, wherein the head mounted display is in a semi-immersive position.

11. The method of claim 1, wherein the lead image is obtained along the line of view of a microscope.

12. The method of claim 1, wherein the lead image is obtained along the line of view of a endoscope.

13. A method for displaying an image of an object, comprising the steps of:
obtaining three dimensional follow image data of the object;
obtaining a real time lead image of the object utilizing a stereo camera;
transforming the three dimensional follow image to correspond to the lead image; and
displaying at least a portion of the transformed three dimensional follow image of the object.

14. The method of claim 13, wherein the stereo camera is at a user's eye level.

15. The method of claim 13, wherein the at least a portion of the transformed three dimensional follow image of the object is displayed on a head mounted display.

16. The method of claim 15, wherein the head mounted display is in a semi-immersive position.

17. The method of claim 13, further comprising automatically determining a desired depth within the object such that the transformed three dimensional follow image of the object represents a two dimensional slice of the object at the desired depth.

18. A method for displaying an image slice of an object, comprising the steps of:
obtaining a real time lead image of the object;
ascertaining a relative three dimensional orientation of the object from the lead image of the object;
calculating desired scan acquistion parameters for a follow image of the object from the lead image of the object;
acquiring the follow image of the object according to the scan acquisition parameters; and
displaying the follow image of the object.

19. The method of claim 18, wherein the lead image of the object is along a user's line of sight to the object.

20. The method of claim 19, wherein the follow image is acquired utilizing an image acquisition machine selected from the group consisting of MRI machines and CT machines.

21. The method of claim 18, wherein the follow image is a two dimensional slice of the object.

22. A head mounted device, comprising:

a camera at a user's eye level that generates images that are input to a computer system; and

a display for displaying two dimensional slices of objects from the computer system that have been transformed to correspond to the line of sight of images from the camera.

23. The device of claim 22, wherein the camera is a stereo camera.

24. The device of claim 22, wherein the camera is a optical localizer.

25. The device of claim 22, wherein the display is semi-immersive.

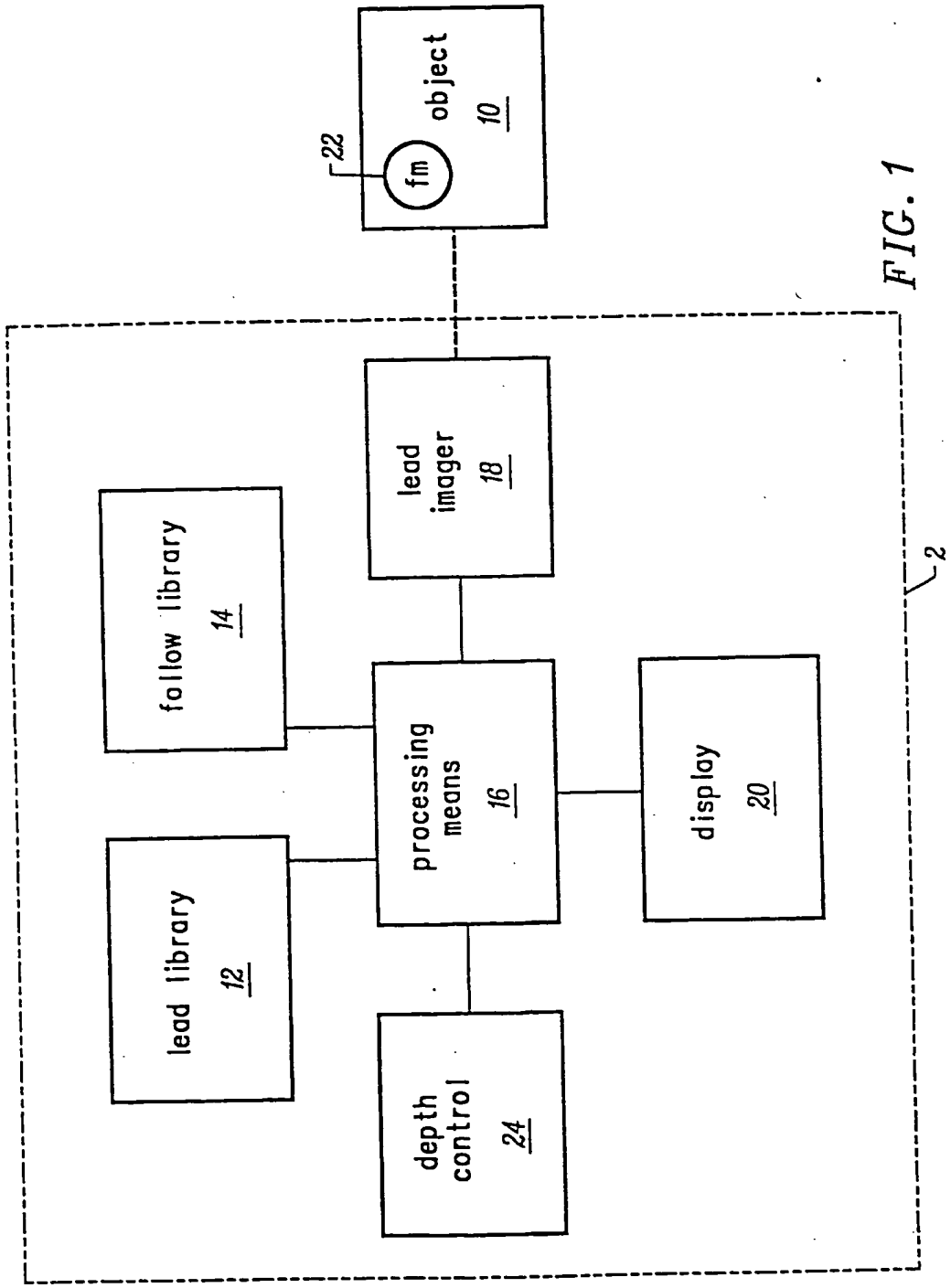

object
10
22
fm
lead
imager
18
follow library
14
processing
means
16
display
20
lead library
12
depth
control
24
2

FIG. 1

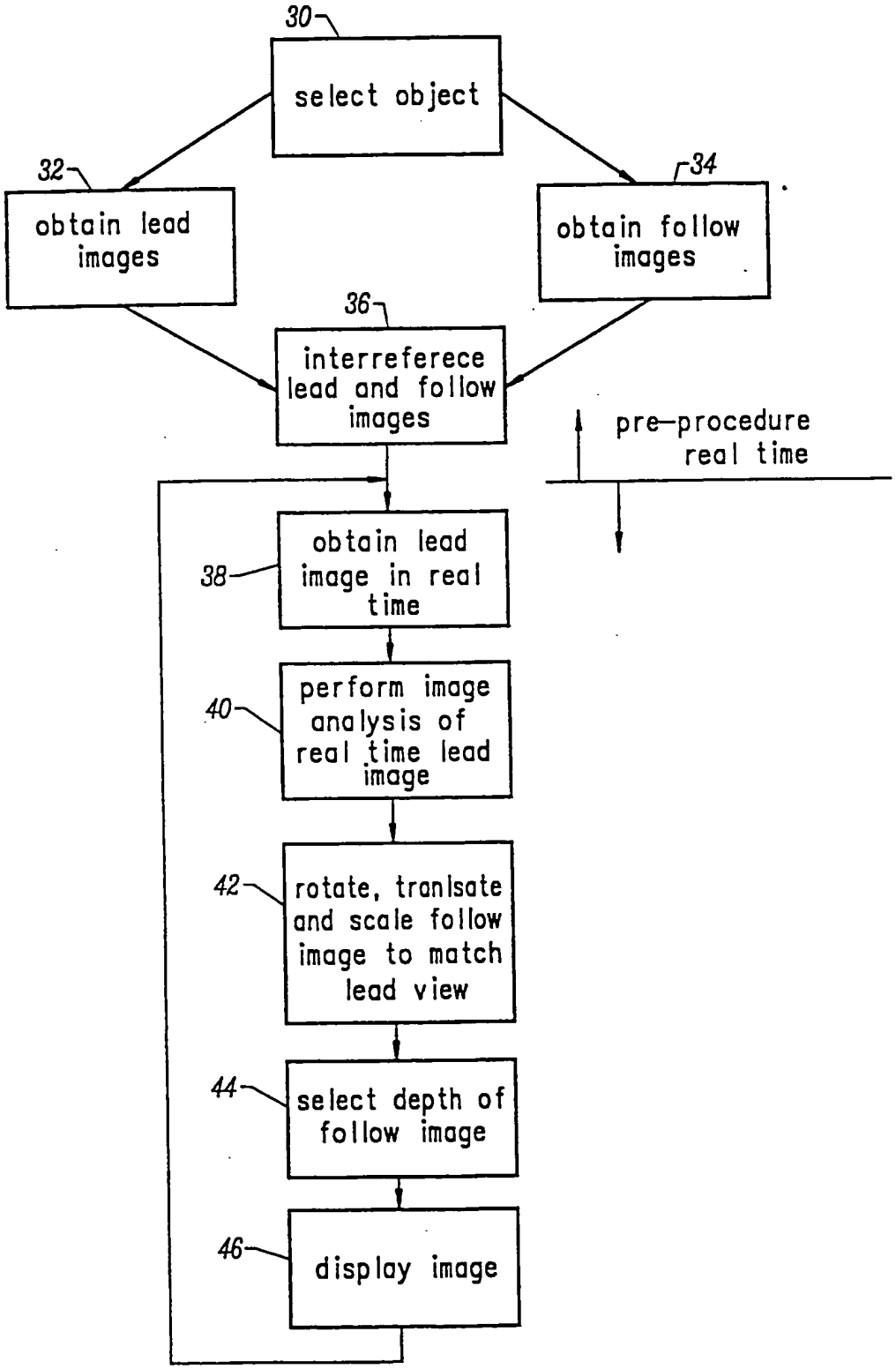

30
select object
32
obtain lead images
34
obtain follow images
36
interreferece lead and follow images
pre-procedure
real time
38
obtain lead image in real time
40
perform image analysis of real time lead image
42
rotate, translsate and scale follow image to match lead view
44
select depth of follow image
46
display image

FIG. 2

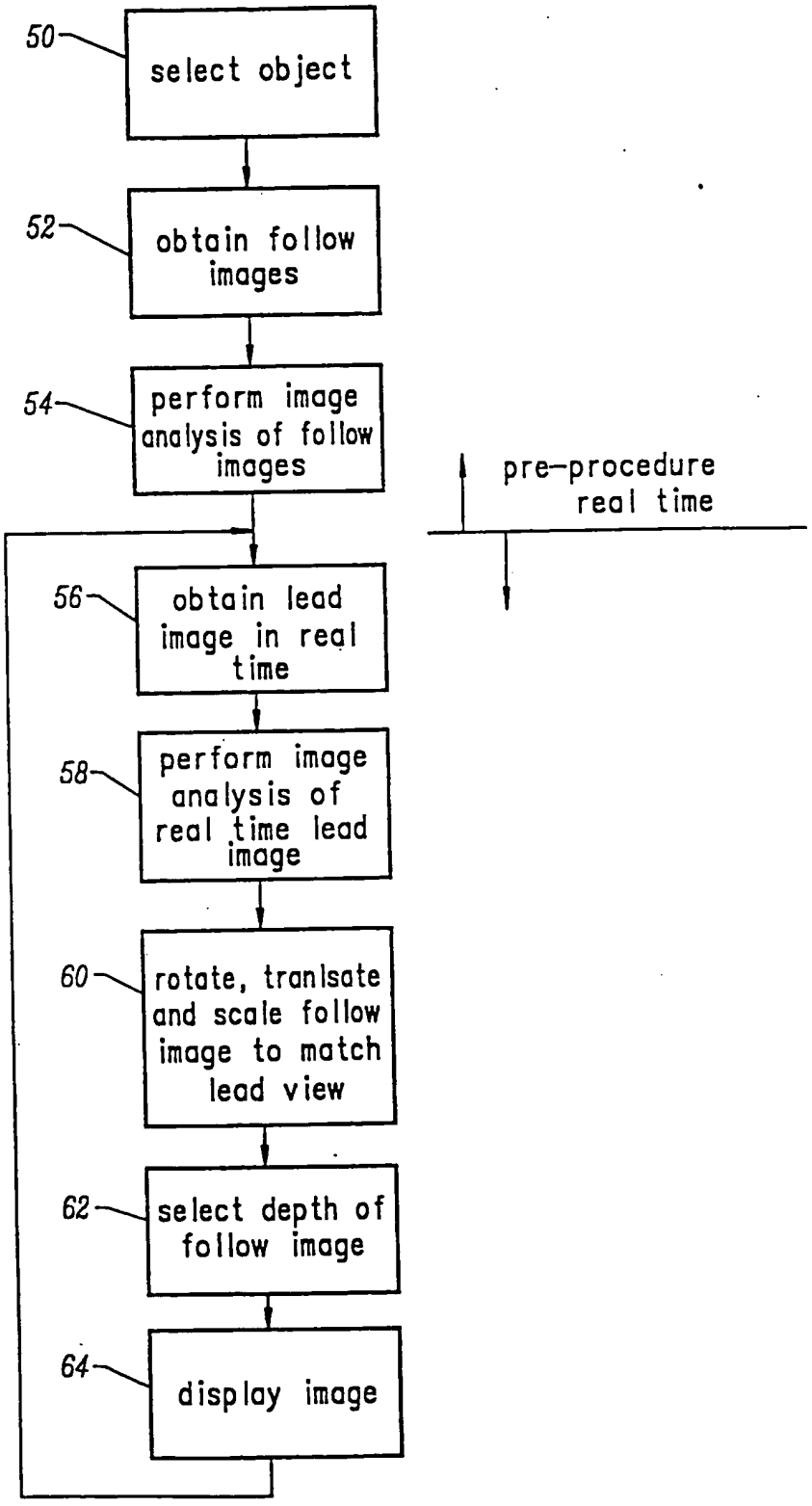

50
select object
52
obtain follow images
54
perform image analysis of follow images
pre-procedure
real time
56
obtain lead image in real time
58
perform image analysis of real time lead image
60
rotate, tranlsate and scale follow image to match lead view
62
select depth of follow image
64
display image

FIG. 3

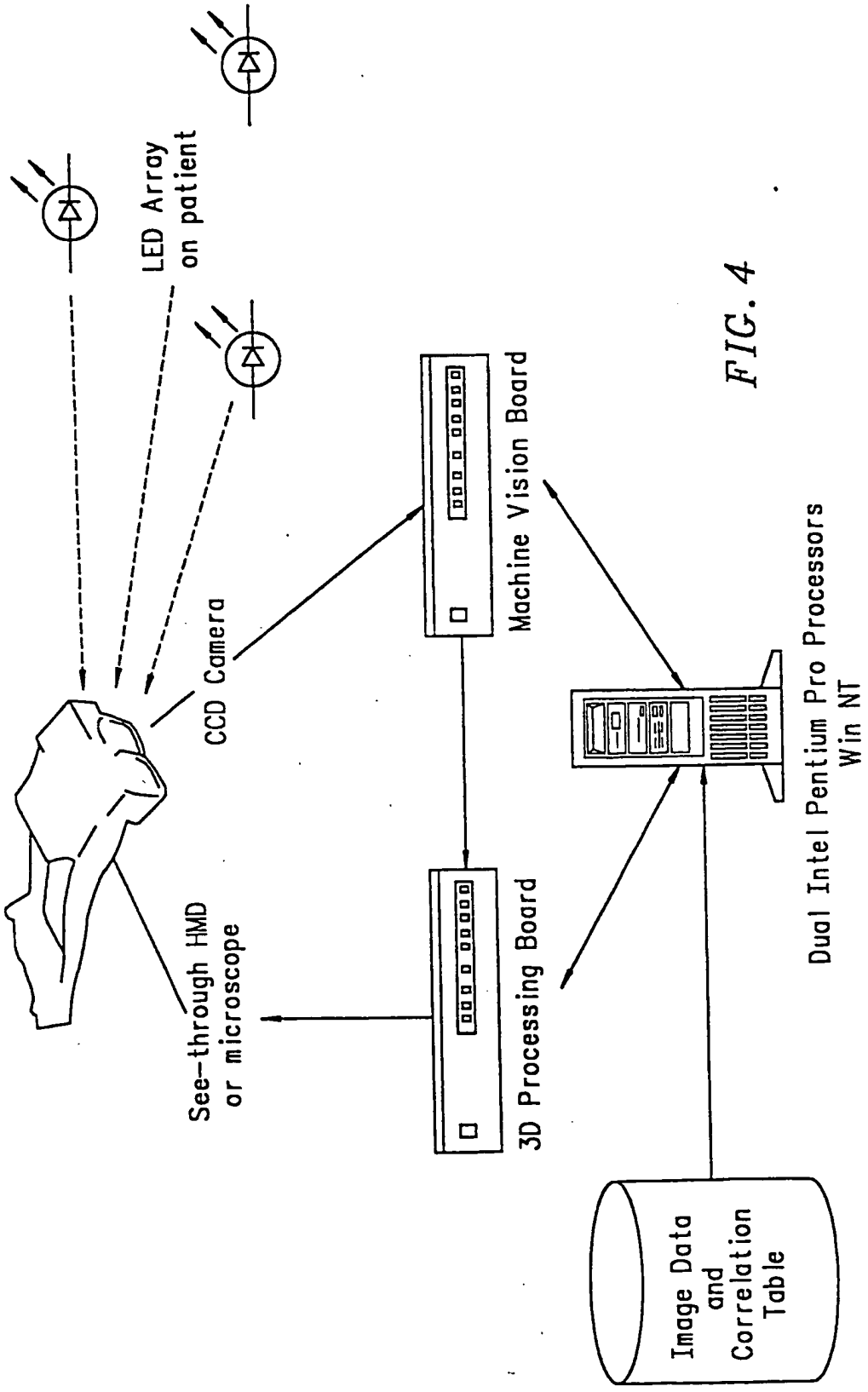

LED Array
on patient
CCD Camera
Machine Vision Board
Dual Intel Pentium Pro Processors
Win NT
See-through HMD
or microscope
3D Processing Board
Image Data
and
Correlation
Table

FIG. 4

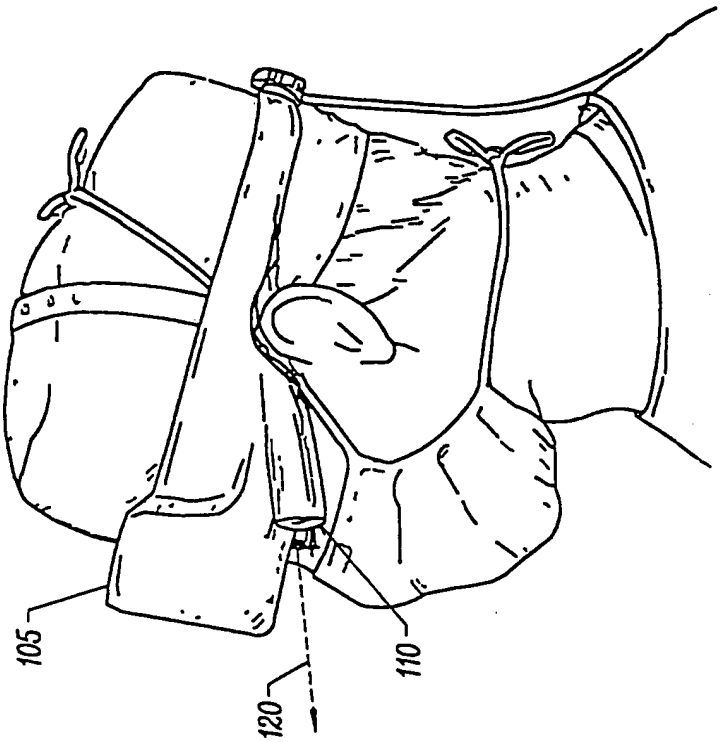


FIG. 5A

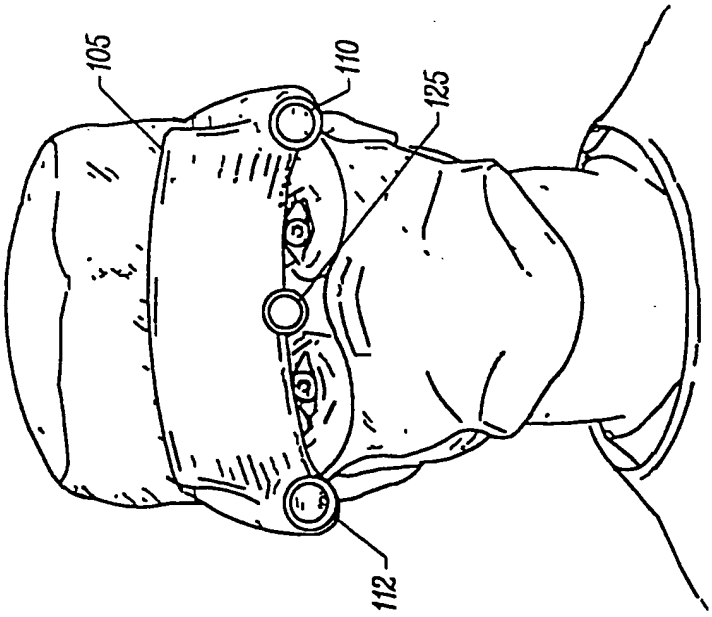


FIG. 5B

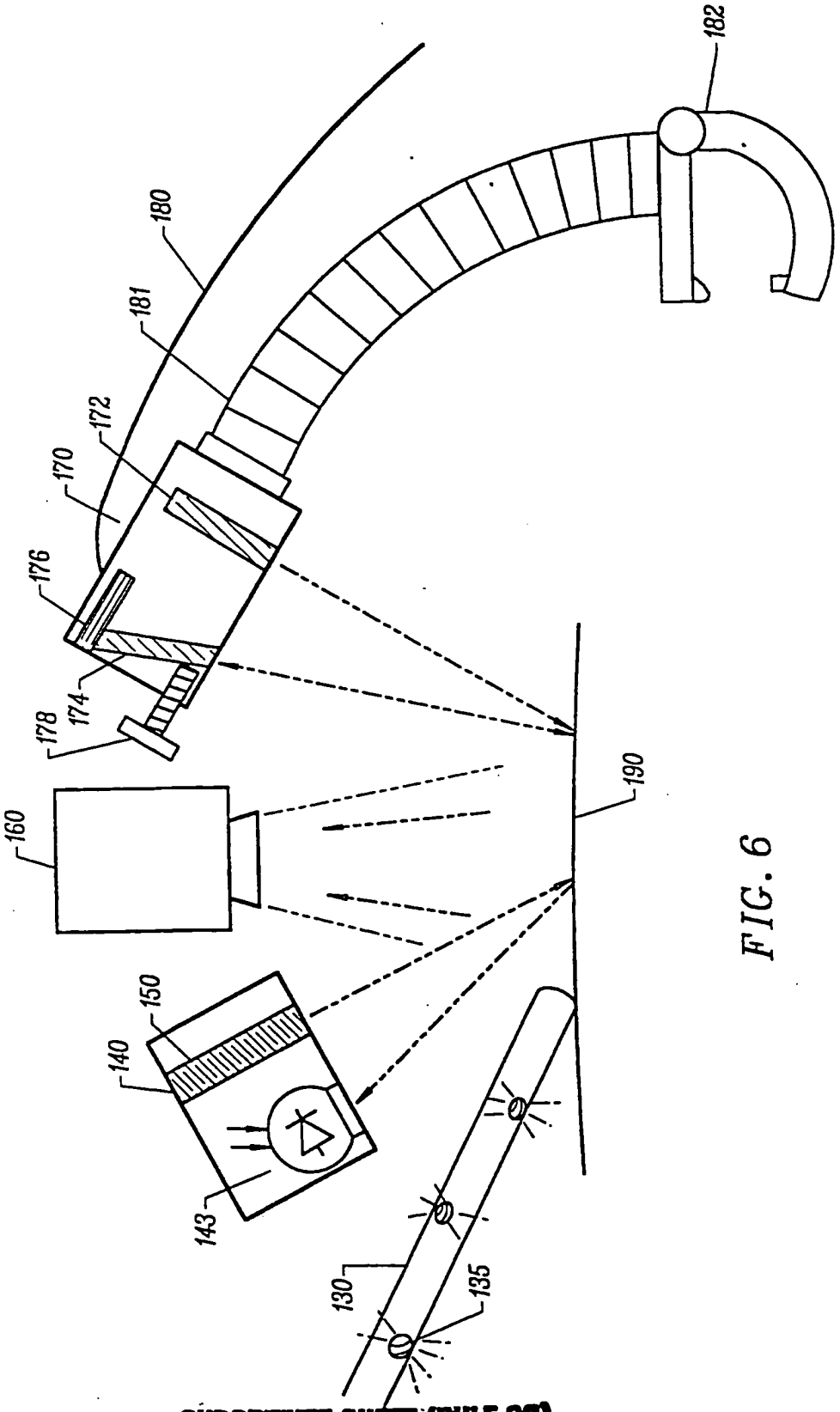


FIG. 6

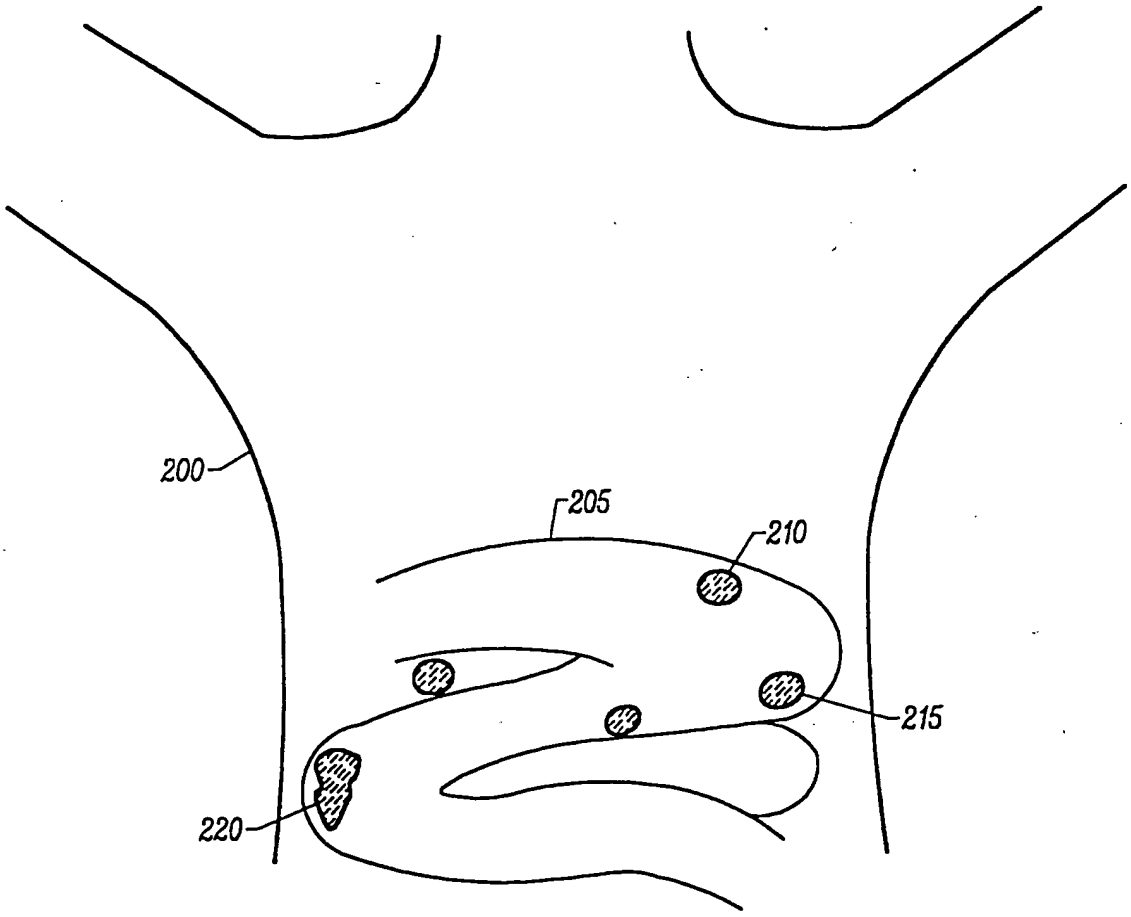


FIG. 7

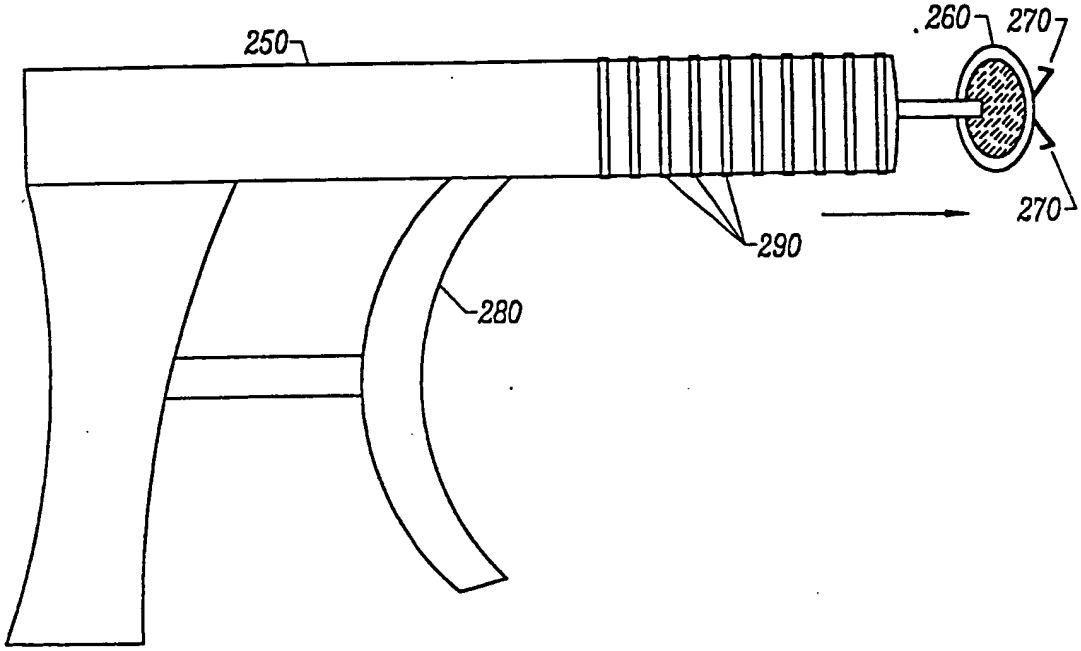


FIG. 8

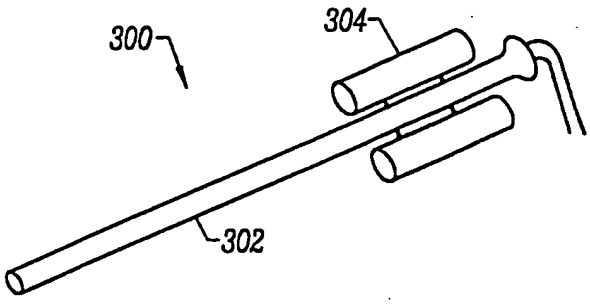


FIG. 9

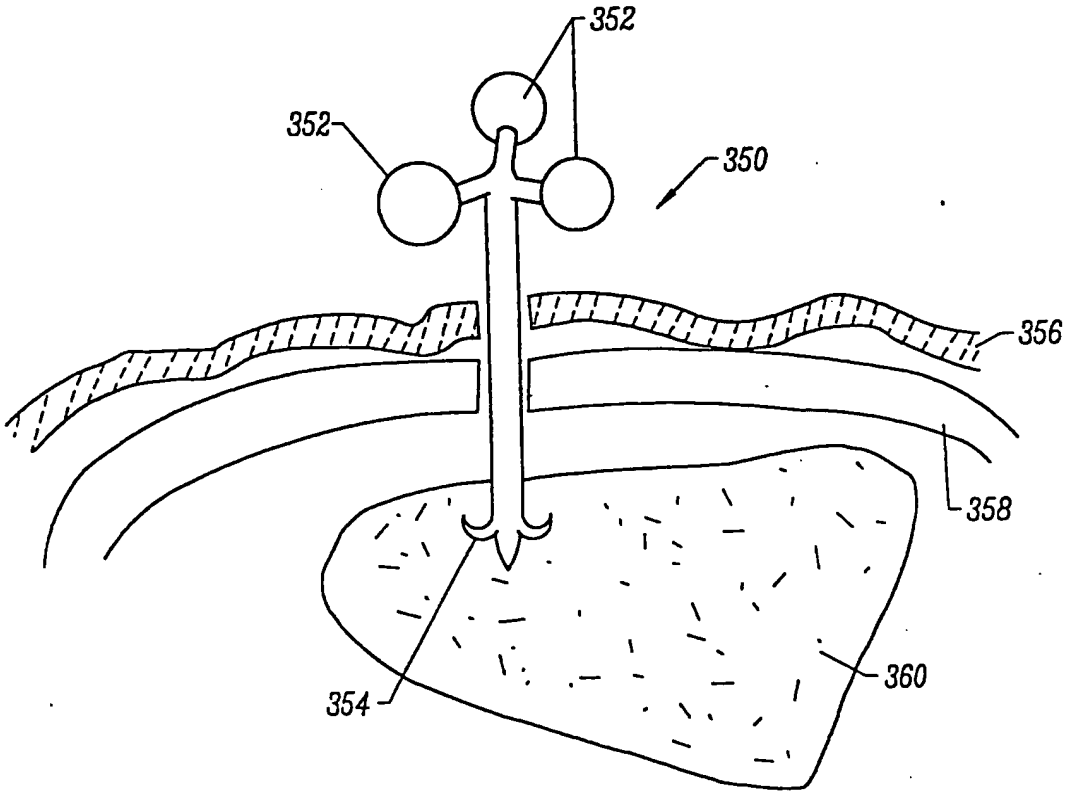


FIG. 10

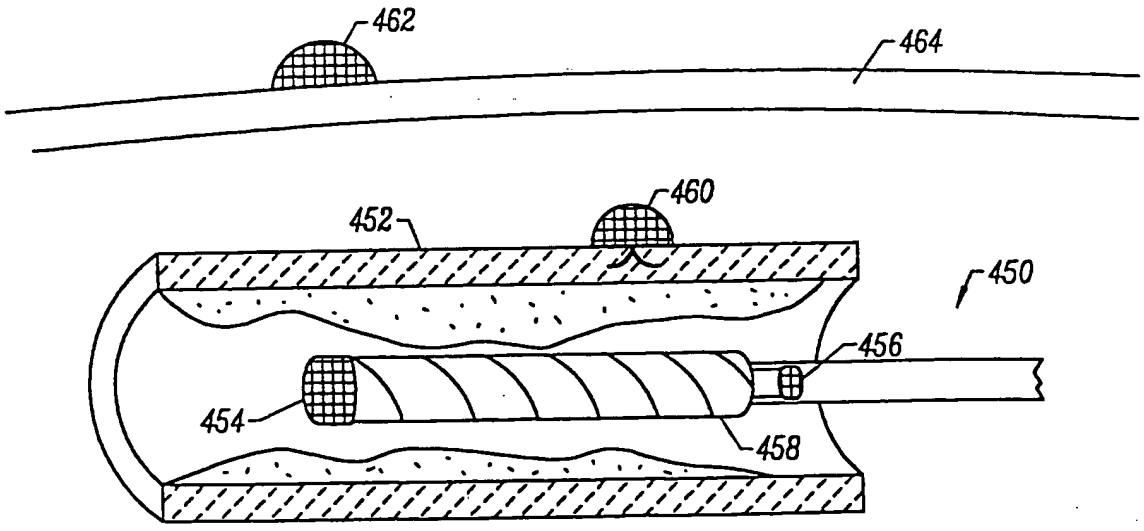


FIG. 12

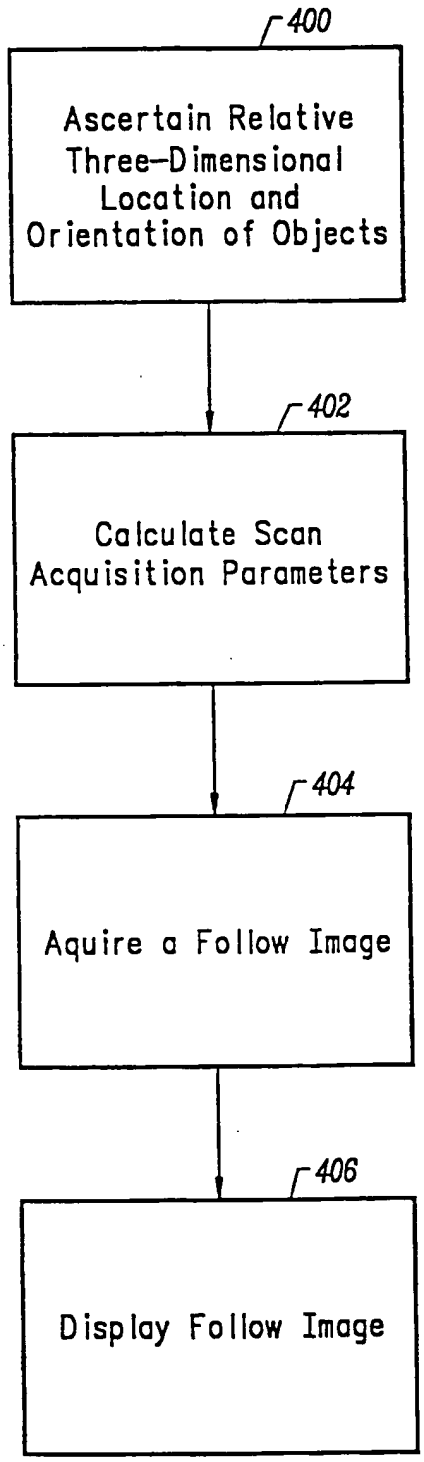

400
Ascertain Relative Three-Dimensional Location and Orientation of Objects
402
Calculate Scan Acquisition Parameters
404
Aquire a Follow Image
406
Display Follow Image

FIG. 11

INTERNATIONAL SEARCH REPORT

International application No.
PCT/US98/04390

A. CLASSIFICATION OF SUBJECT MATTER

IPC(6) :A61B 5/00
US CL :600/425

According to International Patent Classification (IPC) or to both national classification and IPC

B. FIELDS SEARCHED

Minimum documentation searched (classification system followed by classification symbols)

U.S. : 600/407, 410, 425, 426, 436; 378/62, 63, 4, 21; 382/128, 130, 131; 250/370.08, 370.09; 345/115, 121, 139, 427

Documentation searched other than minimum documentation to the extent that such documents are included in the fields searched

NONE

Electronic data base consulted during the international search (name of data base and, where practicable, search terms used)

APS

C. DOCUMENTS CONSIDERED TO BE RELEVANT

Category*	Citation of document, with indication, where appropriate, of the relevant passages	Relevant to claim No.
X	US 5,531,227 A (SCHNEIDER) 02 July 1996, see entire document.	1-25
X	US 5,531,520 A (GRIMSON et al.) 02 July 1996, see entire document.	22-25
A	US 4,343,037 A (BOLTON) 03 August 1982.	1-25
A	US 5,493,595 A (SCHOOLMAN) 20 February 1996.	1-25

☐ Further documents are listed in the continuation of Box C. ☐ See patent family annex.

*	Special categories of cited documents:	"T"	later document published after the international filing date or priority date and not in conflict with the application but cited to understand the principle or theory underlying the invention
"A"	document defining the general state of the art which is not considered to be of particular relevance	"X"	document of particular relevance; the claimed invention cannot be considered novel or cannot be considered to involve an inventive step when the document is taken alone
"E"	earlier document published on or after the international filing date	"Y"	document of particular relevance; the claimed invention cannot be considered to involve an inventive step when the document is combined with one or more other such documents, such combination being obvious to a person skilled in the art
"L"	document which may throw doubts on priority claim(s) or which is cited to establish the publication date of another citation or other special reason (as specified)	"&"	document member of the same patent family
"O"	document referring to an oral disclosure, use, exhibition or other means		
"P"	document published prior to the international filing date but later than the priority date claimed		

Date of the actual completion of the international search

22 JUNE 1998

Date of mailing of the international search report

08 JUL 1998

Name and mailing address of the ISA/US
Commissioner of Patents and Trademarks
Box PCT
Washington, D.C. 20231
Facsimile No. (703) 305-3230

Authorized officer

BRIAN L. CASLER

Telephone No. (703) 308-3552

Form PCT/ISA/210 (second sheet)(July 1992)*

This Page is Inserted by IFW Indexing and Scanning Operations and is not part of the Official Record

BEST AVAILABLE IMAGES

Defective images within this document are accurate representations of the original documents submitted by the applicant.

Defects in the images include but are not limited to the items checked:

- ☐ BLACK BORDERS
- ☐ IMAGE CUT OFF AT TOP, BOTTOM OR SIDES
- ☐ FADED TEXT OR DRAWING
- ☐ BLURRED OR ILLEGIBLE TEXT OR DRAWING
- ☐ SKEWED/SLANTED IMAGES
- ☐ COLOR OR BLACK AND WHITE PHOTOGRAPHS
- ☐ GRAY SCALE DOCUMENTS
- ☐ LINES OR MARKS ON ORIGINAL DOCUMENT
- ☐ REFERENCE(S) OR EXHIBIT(S) SUBMITTED ARE POOR QUALITY
- ☐ OTHER: _____

IMAGES ARE BEST AVAILABLE COPY.
As rescanning these documents will not correct the image problems checked, please do not report these problems to the IFW Image Problem Mailbox.